

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 31, 2004, 13:52:29 ; Search time 3493 Seconds
(without alignments)
7167.840 Million cell updates/sec

Title: US-09-975-856-1
Perfect score: 576
Sequence: 1 ATGACGAGACGACGCTT.....ATGACGAGTAACCTCCCTCG 576

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pin:*
35: em_htg_red:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Length	DB	ID	Description
1	576	100.0	576	6	ARI67456	ARI67456 Sequence
2	576	100.0	576	6	ARI70503	ARI70503 Sequence
3	576	100.0	576	6	ARI82977	ARI82977 Sequence
4	576	100.0	576	6	AR305671	AR305671 Sequence
5	576	100.0	576	6	AX740218	AX740218 Sequence
6	576	100.0	576	6	BD096872	BD096872 Isolated
7	576	100.0	576	6	BD218420	BD218420 SSX gene,
8	576	100.0	576	6	HSU90841	U90841 Homo sapien
9	576	100.0	576	6	BC005325	BC005325 Homo sapi
10	574.4	99.7	1250	9	BC005325	A48452 Sequence 17
11	496.6	86.2	766	6	AX821947	AX821947 Sequence
12	496.6	86.2	766	9	HSSSX2	X86175 H.sapiens m
13	496.6	86.2	931	6	AR025465	AR025465 Sequence
14	496.6	86.2	931	6	AR060380	AR060380 Sequence
15	496.6	86.2	931	6	ARI17874	ARI17874 Sequence
16	496.6	86.2	931	6	ARI67453	ARI67453 Sequence
17	496.6	86.2	931	6	I85576	I85576 Sequence 2
18	496.6	86.2	931	6	AR287596	AR287596 Sequence
19	496.6	86.2	931	6	AR305668	AR305668 Sequence
20	496.6	86.2	931	6	AX114023	AX114023 Sequence
21	496.6	86.2	931	6	AX719099	AX719099 Sequence
22	496.6	86.2	931	6	BD218417	BD218417 SSX gene,
23	496.6	86.2	1309	6	AX331558	AX331558 Sequence
24	496.6	86.2	1309	6	AX331950	AX331950 Sequence
25	496.6	86.2	1309	9	HSXD21LEK	Z49105 H.sapiens H
26	496	86.1	1249	9	HSU90840	U90840 Homo sapien
27	495	85.9	1276	9	BC016957	BC016957 Homo sapi
28	493.4	85.7	1282	9	BC007343	BC007343 Homo sapi
29	492.8	85.6	575	9	S82471	S82471 Homo sapien
30	482.2	83.7	576	6	ARI67457	ARI67457 Sequence
31	482.2	83.7	576	6	ARI70504	ARI70504 Sequence
32	482.2	83.7	576	6	ARI82978	ARI82978 Sequence
33	482.2	83.7	576	6	AR305672	AR305672 Sequence
34	482.2	83.7	576	6	AX740219	AX740219 Sequence
35	482.2	83.7	576	6	BD096873	BD096873 Isolated
36	482.2	83.7	576	6	BD218421	BD218421 SSX gene,
37	481.6	83.6	1281	9	BC001003	BC001003 Homo sapi
38	481.2	83.5	766	6	A48450	A48450 Sequence 15
39	480.6	83.4	576	9	HSU90842	U90842 Homo sapien
40	480	83.3	766	6	ARI67452	ARI67452 Sequence
41	480	83.3	766	6	AR305667	AR305667 Sequence
42	480	83.3	766	6	AX821946	AX821946 Sequence
43	480	83.3	766	6	BD218416	BD218416 SSX gene,
44	480	83.3	766	9	HSSSX1	X86174 H.sapiens m
45	433.2	75.2	1383	9	BC016640	BC016640 Homo sapi

ALIGNMENTS

RESULT 1
ARI67456
LOCUS
DEFINITION Sequence 5 from patent US 6287756.
ACCESSION ARI67456
VERSION ARI67456.1 GI:17903237
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 576)
AUTHORS Tureci,O., Chen,Y.-T., Sahin,U., Gure,A.O., Old,L.J. and
pfreundschuh,M.
TITLE Methods for determining presence of cancer in a sample by
determining expression of an SSX gene

ARI67456 576 bp DNA linear PAT 17-DEC-2001

JOURNAL Patent: US 6287756-A 5 11-SEP-2001;
 FEATURES Location/Qualifiers
 source 1..576
 /organism="unknown"
 /mol_type="unassigned DNA"

ORIGIN

Query Match 100.0%; Score 576; DB 6; Length 576;
 Best Local Similarity 100.0%; Pred. No. 2.6e-164;
 Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAACGAGAGACGAGCCCTTTCACAGGAGACCCAGGATGATGCTCAAAATATCAGAGAAG 60
 |||||
 Db 1 ATGAACGAGAGACGAGCCCTTTCACAGGAGACCCAGGATGATGCTCAAAATATCAGAGAAG 60
 |||||

QY 61 TTACGAAAGGCGCTTCGATGATATGCCAAATATCTCTCTAAGAAAGAGTGGGAAAGATG 120
 |||||
 Db 61 TTACGAAAGGCGCTTCGATGATATGCCAAATATCTCTCTAAGAAAGAGTGGGAAAGATG 120
 |||||

QY 121 AAATCCTCGGAGAAATCGTCTATGTATATGAAGCTAAACTATGAGGTCACTGACTAAA 180
 |||||
 Db 121 AAATCCTCGGAGAAATCGTCTATGTATATGAAGCTAAACTATGAGGTCACTGACTAAA 180
 |||||

QY 181 CTAGGTTTCAAGTCAACCTCCACCTTTTCATGCTAGTAAACGGGCTGCAGACTTCCAC 240
 |||||
 Db 181 CTAGGTTTCAAGTCAACCTCCACCTTTTCATGCTAGTAAACGGGCTGCAGACTTCCAC 240
 |||||

QY 241 GGAATGATTTTGGTAAAGATCGAAACACAGGAATCAGGTTGAACGCTTCAGATGACT 300
 |||||
 Db 241 GGAATGATTTTGGTAAAGATCGAAACACAGGAATCAGGTTGAACGCTTCAGATGACT 300
 |||||

QY 301 TTCCGAGAGCTCCAGAGATCTTCCCGAAGATCATGCCAAGAGCCAGAGAGAGAA 360
 |||||
 Db 301 TTCCGAGAGCTCCAGAGATCTTCCCGAAGATCATGCCAAGAGCCAGAGAGAGAA 360
 |||||

QY 361 AATGGTTTGAAGAGTCCAGAGCACTGCGCCACAAATATGATGGAAACAGCTGTGC 420
 |||||
 Db 361 AATGGTTTGAAGAGTCCAGAGCACTGCGCCACAAATATGATGGAAACAGCTGTGC 420
 |||||

QY 421 CCCCCGGGAAATCCAAAGTACCTTGGAGAGATTAACAGACATCTGGACCCCAAAGGGG 480
 |||||
 Db 421 CCCCCGGGAAATCCAAAGTACCTTGGAGAGATTAACAGACATCTGGACCCCAAAGGGG 480
 |||||

QY 481 AAACATGCTGGACCCACAGACTCGCTGAGAGAAAGCAGCTGGTGGTTTATGAAGATC 540
 |||||
 Db 481 AAACATGCTGGACCCACAGACTCGCTGAGAGAAAGCAGCTGGTGGTTTATGAAGATC 540
 |||||

QY 541 AGCGACCTGAGAGAGATGACGAGTAACCTCCCTCG 576
 |||||
 Db 541 AGCGACCTGAGAGAGATGACGAGTAACCTCCCTCG 576
 |||||

RESULT 2
 AR170503
 LOCUS 576 bp DNA linear PAT 17-DEC-2001
 DEFINITION Sequence 1 from patent US 6291658.
 ACCESSION AR170503
 VERSION AR170503.1 GI:17308462
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unclassified.
 REFERENCE 1 (bases 1 to 576)
 AUTHORS Gure,A.O., Tureci,O., Sahin,U., Tsang,S., Scanlan,M.J., Knuth,A.,
 Pfreundschuh,M., Old,L.J. and Chen,Y.-T.
 TITLE Isolated nucleic acid molecules encoding SSX family members and
 thereof
 JOURNAL Patent: US 6291658-A 1 18-SEP-2001;
 FEATURES Location/Qualifiers
 source 1..576
 /organism="unknown"
 /mol_type="unassigned DNA"

ORIGIN

Query Match 100.0%; Score 576; DB 6; Length 576;
 Best Local Similarity 100.0%; Pred. No. 2.6e-164;
 Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAACGAGAGACGAGCCCTTTCACAGGAGACCCAGGATGATGCTCAAAATATCAGAGAAG 60
 |||||
 Db 1 ATGAACGAGAGACGAGCCCTTTCACAGGAGACCCAGGATGATGCTCAAAATATCAGAGAAG 60
 |||||

QY 61 TTACGAAAGGCGCTTCGATGATATGCCAAATATCTCTCTAAGAAAGAGTGGGAAAGATG 120
 |||||
 Db 61 TTACGAAAGGCGCTTCGATGATATGCCAAATATCTCTCTAAGAAAGAGTGGGAAAGATG 120
 |||||

QY 121 AAATCCTCGGAGAAATCGTCTATGTATATGAAGCTAAACTATGAGGTCACTGACTAAA 180
 |||||
 Db 121 AAATCCTCGGAGAAATCGTCTATGTATATGAAGCTAAACTATGAGGTCACTGACTAAA 180
 |||||

QY 181 CTAGGTTTCAAGTCAACCTCCACCTTTTCATGCTAGTAAACGGGCTGCAGACTTCCAC 240
 |||||
 Db 181 CTAGGTTTCAAGTCAACCTCCACCTTTTCATGCTAGTAAACGGGCTGCAGACTTCCAC 240
 |||||

QY 241 GGAATGATTTTGGTAAAGATCGAAACACAGGAATCAGGTTGAACGCTTCAGATGACT 300
 |||||
 Db 241 GGAATGATTTTGGTAAAGATCGAAACACAGGAATCAGGTTGAACGCTTCAGATGACT 300
 |||||

QY 301 TTCCGAGAGCTCCAGAGATCTTCCCGAAGATCATGCCAAGAGCCAGAGAGAGAA 360
 |||||
 Db 301 TTCCGAGAGCTCCAGAGATCTTCCCGAAGATCATGCCAAGAGCCAGAGAGAGAA 360
 |||||

QY 361 AATGGTTTGAAGAGTCCAGAGCACTGCGCCACAAATATGATGGAAACAGCTGTGC 420
 |||||
 Db 361 AATGGTTTGAAGAGTCCAGAGCACTGCGCCACAAATATGATGGAAACAGCTGTGC 420
 |||||

QY 421 CCCCCGGGAAATCCAAAGTACCTTGGAGAGATTAACAGACATCTGGACCCCAAAGGGG 480
 |||||
 Db 421 CCCCCGGGAAATCCAAAGTACCTTGGAGAGATTAACAGACATCTGGACCCCAAAGGGG 480
 |||||

QY 481 AAACATGCTGGACCCACAGACTCGCTGAGAGAAAGCAGCTGGTGGTTTATGAAGATC 540
 |||||
 Db 481 AAACATGCTGGACCCACAGACTCGCTGAGAGAAAGCAGCTGGTGGTTTATGAAGATC 540
 |||||

QY 541 AGCGACCTGAGAGAGATGACGAGTAACCTCCCTCG 576
 |||||
 Db 541 AGCGACCTGAGAGAGATGACGAGTAACCTCCCTCG 576
 |||||

ORIGIN

Query Match 100.0%; Score 576; DB 6; Length 576;
 Best Local Similarity 100.0%; Pred. No. 2.6e-164;
 Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAACGAGAGACGAGCCCTTTCACAGGAGACCCAGGATGATGCTCAAAATATCAGAGAAG 60
 |||||

121 AAATCCTCGGAGAAATCGTCTATGTATATGAGCTAAACTATGAGTCACTATAA 180
121 AAATCCTCGGAGAAATCGTCTATGTATATGAGCTAAACTATGAGTCACTATAA 180
181 CTAGGTTTCAAGGTCACCTCCACCTTTCATCGTAGTAAACGGGCTGAGACTTCCAC 240
181 CTAGGTTTCAAGGTCACCTCCACCTTTCATCGTAGTAAACGGGCTGAGACTTCCAC 240
241 GGGATGATTTTGGTAAACGATCGAAACACACAGGAATCAGGTTTGAACGCTCCTCAGATGACT 300
241 GGGATGATTTTGGTAAACGATCGAAACACACAGGAATCAGGTTTGAACGCTCCTCAGATGACT 300
301 TTGGGAGCCTTCAGAGAAATCTTCCGAGATCATGCCAGAGACAGGAGAGAGAA 360
301 TTGGGAGCCTTCAGAGAAATCTTCCGAGATCATGCCAGAGACAGGAGAGAGAA 360
361 AATGTTTGAAGGAGTCCAGAGGATCTCGCCACCAAAATGATGGGAAAACAGCTGTGC 420
361 AATGTTTGAAGGAGTCCAGAGGATCTCGCCACCAAAATGATGGGAAAACAGCTGTGC 420
421 CCCCCGGGAAATCCAAATGATCTTGGAGAGATTAACAGACATCTGGACCCCAAAAGGGGG 480
421 CCCCCGGGAAATCCAAATGATCTTGGAGAGATTAACAGACATCTGGACCCCAAAAGGGGG 480
481 AAACATGCTCGGACCCAGAGCTCGTGGAGAGAAAGCAGCTGGTGGTTTATGAGAGATC 540
481 AAACATGCTCGGACCCAGAGCTCGTGGAGAGAAAGCAGCTGGTGGTTTATGAGAGATC 540
541 AGCGACCTCGAGGAAAGTACGAGTAATCCCTCCG 576
541 AGCGACCTCGAGGAAAGTACGAGTAATCCCTCCG 576

RESULT 5
AX740218 576 bp mRNA linear PAT 08-MAY-2003
LOCUS Sequence 1 from Patent.EP1300463.
DEFINITION AX740218
ACCESSION AX740218
VERSION AX740218.1 GI:30519358
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1
AUTHORS Gure,A.O., Tureci,O., Sahin,U., Tsang,S., Scanlan,M.J., Knuth,A.,
Pfreundschuh,M., Old,L.J. and Chen,Y.T.
TITLE Isolated nucleic acid molecules encoding SSX family members and
uses thereof
JOURNAL Patent: EP 1300463-A 1 09-APR-2003;
LUDWIG INSTITUTE FOR CANCER RESEARCH (US) ; MEMORIAL
SLOAN-KETTERING CANCER CENTER (US) ; CORNELL RESEARCH FOUNDATION
(US)

FEATURES
source Location/Qualifiers
1..576 /organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
gene 1..576
CDS 1..567
/gene="SSX4"
/codon_start=1
/protein_id="CAD90570.1"
/db_xref="GI:30519359"
/translation="NMDDAFARRPRDDAQISEKLFKAFDDIAKYFKSEKWEKMSSE
KIVVYVKNLYEVMKLGFKVTLPPMRSKRAADPHGNDFGNDRNHRNOVERPQMTFG
SLORIPKINPKPAEENGLKEVPEASGPQNDGKQLCPFGNPSTLEKINKTSGPKRG
KHAWHLRLRKQLVVEIEISDFEEDDE"

ORIGIN
Query Match 100.0%; Score 576; DB 6; Length 576;

1 ATGAACGGAGACGACGCCCTTTCAGAGAGACCCAGGATGATGCTCAATATATCAGAGAG 60
61 TTACGAAAGCCCTTCGATGATATGCAAAATATCTTCTAAGAAAGAGTGGGAAAGATG 120
61 TTACGAAAGCCCTTCGATGATATGCAAAATATCTTCTAAGAAAGAGTGGGAAAGATG 120
121 AAATCCTCGGAGAAATCGTCTATGTATATGAGCTAAACTATGAGTCACTATAA 180
121 AAATCCTCGGAGAAATCGTCTATGTATATGAGCTAAACTATGAGTCACTATAA 180
181 CTAGGTTTCAAGGTCACCTCCACCTTTCATCGTAGTAAACGGGCTGAGACTTCCAC 240
181 CTAGGTTTCAAGGTCACCTCCACCTTTCATCGTAGTAAACGGGCTGAGACTTCCAC 240
241 GGGATGATTTTGGTAAACGATCGAAACACAGGAATCAGGTTTGAACGCTCCTCAGATGACT 300
241 GGGATGATTTTGGTAAACGATCGAAACACAGGAATCAGGTTTGAACGCTCCTCAGATGACT 300
301 TTGGGAGCCTTCAGAGAAATCTTCCGAGATCATGCCAGAGACAGGAGAGAGAA 360
301 TTGGGAGCCTTCAGAGAAATCTTCCGAGATCATGCCAGAGACAGGAGAGAGAA 360
361 AATGTTTGAAGGAGTCCAGAGGATCTCGCCACCAAAATGATGGGAAAACAGCTGTGC 420
361 AATGTTTGAAGGAGTCCAGAGGATCTCGCCACCAAAATGATGGGAAAACAGCTGTGC 420
421 CCCCCGGGAAATCCAAATGATCTTGGAGAGATTAACAGACATCTGGACCCCAAAAGGGGG 480
421 CCCCCGGGAAATCCAAATGATCTTGGAGAGATTAACAGACATCTGGACCCCAAAAGGGGG 480
481 AAACATGCTCGGACCCAGAGCTCGTGGAGAGAAAGCAGCTGGTGGTTTATGAGAGATC 540
481 AAACATGCTCGGACCCAGAGCTCGTGGAGAGAAAGCAGCTGGTGGTTTATGAGAGATC 540
541 AGCGACCTCGAGGAAAGTACGAGTAATCCCTCCG 576
541 AGCGACCTCGAGGAAAGTACGAGTAATCCCTCCG 576

RESULT 4
AR305671 576 bp DNA linear PAT 12-JUN-2003
LOCUS Sequence 5 from patent US 6548064.
DEFINITION AR305671
ACCESSION AR305671
VERSION AR305671.1 GI:31695155
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 576)
AUTHORS Tureci,O., Sahin,U., Pfreundschuh,M., Rammensee,H.G. and
Stevanovic,S.
TITLE Isolated peptides consisting of amino acid sequences found in SSX
or NY-ESO-1 molecules, which bind to HLA molecules
JOURNAL Patent: US 6548064-A 5 15-APR-2003;
Location/Qualifiers
source 1..576
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 100.0%; Score 576; DB 6; Length 576;
Best Local Similarity 100.0%; Pred. No. 2.6e-164;
Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAACGGAGACGACGCCCTTTCAGAGAGACCCAGGATGATGCTCAATATATCAGAGAG 60
DB 1 ATGAACGGAGACGACGCCCTTTCAGAGAGACCCAGGATGATGCTCAATATATCAGAGAG 60
QY 1 ATGAACGGAGACGACGCCCTTTCAGAGAGACCCAGGATGATGCTCAATATATCAGAGAG 120
DB 1 ATGAACGGAGACGACGCCCTTTCAGAGAGACCCAGGATGATGCTCAATATATCAGAGAG 120
QY 1 TTACGAAAGCCCTTCGATGATATGCAAAATATCTTCTAAGAAAGAGTGGGAAAGATG 180
DB 1 TTACGAAAGCCCTTCGATGATATGCAAAATATCTTCTAAGAAAGAGTGGGAAAGATG 180
QY 1 TTACGAAAGCCCTTCGATGATATGCAAAATATCTTCTAAGAAAGAGTGGGAAAGATG 240
DB 1 TTACGAAAGCCCTTCGATGATATGCAAAATATCTTCTAAGAAAGAGTGGGAAAGATG 240

Best Local Similarity 100.0%; Pred. No. 2.6e-164; Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAACGAGACGACGCTTTGCAAGGAGACCCAGGATGATCTCAATATATCAGAGAAG 60
 Db 1 ATGAACGAGACGACGCTTTGCAAGGAGACCCAGGATGATCTCAATATATCAGAGAAG 60
 QY 61 TTACGAAGGCTTCGATGATATTCGCAATATCTCTTAAGAAAGAGTGGAAAGATG 120
 Db 61 TTACGAAGGCTTCGATGATATTCGCAATATCTCTTAAGAAAGAGTGGAAAGATG 120
 QY 121 AAATCCTCGGAGAAATCGTCTATGTATATGAAGCTAACTATGAGTCTATCACTAA 180
 Db 121 AAATCCTCGGAGAAATCGTCTATGTATATGAAGCTAACTATGAGTCTATCACTAA 180
 QY 181 CTAGTTTCAAGGTCAACCTCCACCTTCATGCTAGTAAACGGGCTGAGACTCCAC 240
 Db 181 CTAGTTTCAAGGTCAACCTCCACCTTCATGCTAGTAAACGGGCTGAGACTCCAC 240
 QY 241 GGGATGATTTTGGTAAAGTCAAGAACCAAGGATCAGTTGAACCTCTCAGATGACT 300
 Db 241 GGGATGATTTTGGTAAAGTCAAGAACCAAGGATCAGTTGAACCTCTCAGATGACT 300
 QY 301 TTGGCAGCTCCAGAGAAATCTTCCGAAGATCATGCCAAGAGCCAGCAGAGAGAA 360
 Db 301 TTGGCAGCTCCAGAGAAATCTTCCGAAGATCATGCCAAGAGCCAGCAGAGAGAA 360
 QY 361 AATGTTTGAAGGAGTCCAGAGATCTCGGCCCAACAAATGATGGAAACAGCTGTC 420
 Db 361 AATGTTTGAAGGAGTCCAGAGATCTCGGCCCAACAAATGATGGAAACAGCTGTC 420
 QY 421 CCCCAGGAAATCCAAAGTACCTTGGAGAAATTAACAGAGATCTGAGACCATCGA 480
 Db 421 CCCCAGGAAATCCAAAGTACCTTGGAGAAATTAACAGAGATCTGAGACCATCGA 480
 QY 481 AAACATGCTGAGACCCAGAGATCTGCTGAGAGAAAGAGCTGGTGTATGAAGATC 540
 Db 481 AAACATGCTGAGACCCAGAGATCTGCTGAGAGAAAGAGCTGGTGTATGAAGATC 540
 QY 541 AGCGACCTCGAGAAAGATGACGAGTAACTCCCTCG 576
 Db 541 AGCGACCTCGAGAAAGATGACGAGTAACTCCCTCG 576

RESULT 6
 BD096872
 LOCUS
 DEFINITION
 Isolated nucleic acid molecules encoding SSX family members and uses thereof.
 ACCESSION
 BD096872
 VERSION
 JP 2001527408-A/1.
 KEYWORDS
 unclassified
 SOURCE
 unclassified
 ORGANISM
 unclassified
 REFERENCE
 1 (bases 1 to 576)
 Gure,A.O., Tureci,O., Sahin,U., Teang,S., Scanlan,M.J., Knuth,A., Pire,A.O., Tureci,M., Old,L.J. and Chen,Y.T.
 AUTHORS
 Isolated nucleic acid molecules encoding SSX family members and uses thereof
 JOURNAL
 Patent: JP 2001527408-A 1 25-DEC-2001;
 LUDWIG INSTITUTE FOR CANCER RESEARCH, MEMORIAL SLOAN KETTERING
 CANCER CENTER, CORNELL RESEARCH FOUNDATION
 COMMENT
 OS Unidentified
 PN JP 2001527408-A/1
 PD 25-DEC-2001
 PF 25-FEB-1998 JP 1998548050
 PR 05-MAY-1997 US 08/851138
 PI ALI O GURE, OZLEM TURECI, UGUR SAHIN, SOLAM TSANG, MATTHEW J PI
 SCANLAN,
 PI ALEXANDER KNUTH, MICHAEL PEREUNDSCHUH, LLOYD J OLD, YAO TSENG PI
 CHEN
 PC C12N5/10, C12N15/12, C12P21/02, C12Q1/68

CC Strandedness: Single;
 CC Topology: Linear;
 CC Isolated nucleic acid molecules encoding SSX family members
 CC thereof
 CC Key Location/Qualifiers
 FH source 1..576
 FT /organism="Unidentified",
 FT Location/Qualifiers
 1..576
 /organism="unidentified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"

FEATURES
 source

ORIGIN
 Query Match 100.0%; Score 576; DB 6; Length 576;
 Best Local Similarity 100.0%; Pred. No. 2.6e-164;
 Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAACGAGACGACGCTTTGCAAGGAGACCCAGGATGATCTCAATATATCAGAGAAG 60
 Db 1 ATGAACGAGACGACGCTTTGCAAGGAGACCCAGGATGATCTCAATATATCAGAGAAG 60
 QY 61 TTACGAAGGCTTCGATGATATTCGCAATATCTCTTAAGAAAGAGTGGAAAGATG 120
 Db 61 TTACGAAGGCTTCGATGATATTCGCAATATCTCTTAAGAAAGAGTGGAAAGATG 120
 QY 121 AAATCCTCGGAGAAATCGTCTATGTATATGAAGCTAACTATGAGTCTATCACTAA 180
 Db 121 AAATCCTCGGAGAAATCGTCTATGTATATGAAGCTAACTATGAGTCTATCACTAA 180
 QY 181 CTAGTTTCAAGGTCAACCTCCACCTTCATGCTAGTAAACGGGCTGAGACTCCAC 240
 Db 181 CTAGTTTCAAGGTCAACCTCCACCTTCATGCTAGTAAACGGGCTGAGACTCCAC 240
 QY 241 GGGATGATTTTGGTAAAGTCAAGAACCAAGGATCAGTTGAACCTCTCAGATGACT 300
 Db 241 GGGATGATTTTGGTAAAGTCAAGAACCAAGGATCAGTTGAACCTCTCAGATGACT 300
 QY 301 TTGGCAGCTCCAGAGAAATCTTCCGAAGATCATGCCAAGAGCCAGCAGAGAGAA 360
 Db 301 TTGGCAGCTCCAGAGAAATCTTCCGAAGATCATGCCAAGAGCCAGCAGAGAGAA 360
 QY 361 AATGTTTGAAGGAGTCCAGAGATCTCGGCCCAACAAATGATGGAAACAGCTGTC 420
 Db 361 AATGTTTGAAGGAGTCCAGAGATCTCGGCCCAACAAATGATGGAAACAGCTGTC 420
 QY 421 CCCCAGGAAATCCAAAGTACCTTGGAGAAATTAACAGAGATCTGAGACCATCGA 480
 Db 421 CCCCAGGAAATCCAAAGTACCTTGGAGAAATTAACAGAGATCTGAGACCATCGA 480
 QY 481 AAACATGCTGAGACCCAGAGATCTGCTGAGAGAAAGAGCTGGTGTATGAAGATC 540
 Db 481 AAACATGCTGAGACCCAGAGATCTGCTGAGAGAAAGAGCTGGTGTATGAAGATC 540
 QY 541 AGCGACCTCGAGAAAGATGACGAGTAACTCCCTCG 576
 Db 541 AGCGACCTCGAGAAAGATGACGAGTAACTCCCTCG 576

RESULT 7

BD218420

LOCUS

DEFINITION

SSX gene, method of determining the occurrence of cancer in sample

by determining the expression of peptides originating in the SSX

gene and NY-ESO-1 gene and utilization thereof.

ACCSSION

BD218420

VERSION

JP 2002519013-A/5.

KEYWORDS

Homo sapiens (human)

SOURCE

Homo sapiens

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

BD218420 576 bp DNA linear PAT 17-JUL-2003
 SSX gene, method of determining the occurrence of cancer in sample
 by determining the expression of peptides originating in the SSX
 gene and NY-ESO-1 gene and utilization thereof.

BD218420
 BD218420.1 GI:33028190
 JP 2002519013-A/5.
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 576)
Tureci, O., Sahin, U., Pfeundschtuh, M., Rammensee, G., Stevanovic, S.,
Chen, Y.T., Gure, A. and Old, L.J., the occurrence of cancer in sample
SSX gene, method of determining the expression of peptides originating in the SSX
gene and NY-ESO-1 gene and utilization thereof
Patent: JP 2002519013-A 5 02-JUL-2002;
LUDWIG INSTITUTE FOR CANCER RESEARCH
OS Homo sapiens (human)
PN JP 2002519013-A/5
PD 02-JUN-2002 JP 2000557145
PF 25-JUN-1999 JP 2000557145
PR 26-JUN-1998 US 09/105839
PI OZLEM TURECI, UGUR SAHIN, MICHAEL PFEUNDSCHUH, GEORG RAMMENSEE,
STEFAN STEVANOVIC, YAO TSENG CHEN, ALI GURE, ILOYD J OLD PC
C12N15/09, A61K38/00, C07H21/04, C12N5/10, C12P21/04, C12Q1/68, PC
G01N33/48
PC G01N33/53, G01N33/574, C12N15/00, A61K37/02, C12N5/00 CC SSX
gene, method of determining the occurrence of cancer in CC
sample by
CC determining the expression of peptides originating in the SSX
gene and
CC NY-ESO-1 gene and utilization thereof
FH Key Location/Qualifiers
FT source 1..576
FT Location/Qualifiers
FEATURES
source
1..576
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 576; DB 6; Length 576;
Best Local Similarity 100.0%; Pred. No. 2.6e-164;
Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGAACGAGACGACGCTTTGCAAGGAGACCCAGGATGATGCTCAATATCAGAGAAG 60
Db 1 ATGAACGAGACGACGCTTTGCAAGGAGACCCAGGATGATGCTCAATATCAGAGAAG 60
Qy 61 TTACGAAAGGCGCTTCGATGATATTGCCAATATCTTCTAAGAAAGAGTGGGAAAGATG 120
Db 61 TTACGAAAGGCGCTTCGATGATATTGCCAATATCTTCTAAGAAAGAGTGGGAAAGATG 120
Qy 121 AAATCCTCGGAGAAAACGCTATGTGTATAGAGTAACTAATGAGGTCATGACTTAA 180
Db 121 AAATCCTCGGAGAAAACGCTATGTGTATAGAGTAACTAATGAGGTCATGACTTAA 180
Qy 181 CTAGGTTTCAAGGTCACCTCCACCTTTCATGCGTAGTAAACGGGCTGCAGACTTCCAC 240
Db 181 CTAGGTTTCAAGGTCACCTCCACCTTTCATGCGTAGTAAACGGGCTGCAGACTTCCAC 240
Qy 241 GGGAAATGATTTTGTAGCATCGAAACACAGGAATCAGGTTGAACGCTCCTCAGATGACT 300
Db 241 GGGAAATGATTTTGTAGCATCGAAACACAGGAATCAGGTTGAACGCTCCTCAGATGACT 300
Qy 301 TTCCGACGCTCCAGAGAAATCTTCCGAGAGATCATGCCAAGAGCCAGCAGAGAGAA 360
Db 301 TTCCGACGCTCCAGAGAAATCTTCCGAGAGATCATGCCAAGAGCCAGCAGAGAGAA 360
Qy 361 AATGTTTGAAGGAAGTCCAGAGCATCTGCGCCACAAAATGATGGAAACAGCTGTGC 420
Db 361 AATGTTTGAAGGAAGTCCAGAGCATCTGCGCCACAAAATGATGGAAACAGCTGTGC 420
Qy 421 CCCCAGGAAATCCAGTACCTTGGAGAGATTAACAGACATCTGGACCCAAAGGGGG 480
Db 421 CCCCAGGAAATCCAGTACCTTGGAGAGATTAACAGACATCTGGACCCAAAGGGGG 480
Qy 481 AAACATGCTCGGACCCACAGACTGCTGAGAGAAAAGCAGCTGGTGGTTTATGAAGATC 540
Db 481 AAACATGCTCGGACCCACAGACTGCTGAGAGAAAAGCAGCTGGTGGTTTATGAAGATC 540

Qy 541 AGGACCCCTGAGGAGATGACGAGTAACCTCCCTCG 576
Db 541 AGGACCCCTGAGGAGATGACGAGTAACCTCCCTCG 576
RESULT 8
HSU0841
LOCUS Homo sapiens SSX4 (SSX4) mRNA, complete cds. PRI 18-MAR-1998
DEFINITION Homo sapiens SSX4 (SSX4) mRNA, complete cds.
ACCESSION U90841
VERSION U90841.1 GI:2952022
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 576)
Gure, A.O., Tureci, O., Sahin, U., Tsang, S., Scanlan, M.J., Jager, E.,
Knuth, A., Pfeundschtuh, M., Old, L.J. and Chen, Y.T.
SSX: a multigene family with several members transcribed in normal
testis and human cancer.
Int. J. Cancer 72 (6), 965-971 (1997)
JOURNAL 98021352
MEDLINE 9378559
PUBMED
REFERENCE 2 (bases 1 to 576)
Gure, A.O., Tureci, O., Sahin, U., Tsang, S., Scanlan, M., Knuth, A.,
Pfeundschtuh, M., Old, L.J. and Chen, Y.T.
Direct Submission
Submitted (05-FEB-1997) Pathology C-320, Cornell University, 1300
York Ave., New York City, NY 10021, USA
LOCATION/Qualifiers
1..576
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..576
/gene="SSX4"
1..567
/gene="SSX4"
/note="putative translation product; Kruppel-associated
box containing SSX gene"
/codon_start=1
/product="SSX4"
/protein_id="AAC05820.1"
/db_xref="GI:2952023"
translation="MNGDDAFARRPRDDAQISEKRAFDDIAKFKSEKWEKMSSE
KIVVYVLMNYEVMTKGFVTLPPFMRSKKAADFGNDFGNDRHNRQVPRQMTFG
SLQRFPPKIMPKPAEENGKLEKVEASGPNQDKQLCPGPNPSTLEKINKTSGPKRG
KAWTHRLRERKQLVYVEISDPEDDE"
1..18
/gene="SSX4"
/note="corresponds to the 5' PCR primer that was used to
isolate this gene and is identical to SSX2, GenBank
Accession Number X86175; therefore, the SSX4 gene sequence
might contain differences within this span"
primer_bind
Query Match 100.0%; Score 576; DB 9; Length 576;
Best Local Similarity 100.0%; Pred. No. 2.6e-164;
Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGAACGAGACGACGCTTTGCAAGGAGACCCAGGATGATGCTCAATATCAGAGAAG 60
Db 1 ATGAACGAGACGACGCTTTGCAAGGAGACCCAGGATGATGCTCAATATCAGAGAAG 60
Qy 61 TTACGAAAGGCGCTTCGATGATATTGCCAATATCTTCTTAAAGAGAGTGGGAAAGATG 120
Db 61 TTACGAAAGGCGCTTCGATGATATTGCCAATATCTTCTTAAAGAGAGTGGGAAAGATG 120
Qy 121 AAATCCTCGGAGAAAACGCTATGTGTATAGAGTAACTAATGAGGTCATGACTTAA 180
Db 121 AAATCCTCGGAGAAAACGCTATGTGTATAGAGTAACTAATGAGGTCATGACTTAA 180
Qy 181 CTAGGTTTCAAGGTCACCTCCACCTTTCATGCGTAGTAAACGGGCTGCAGACTTCCAC 240
Db 181 CTAGGTTTCAAGGTCACCTCCACCTTTCATGCGTAGTAAACGGGCTGCAGACTTCCAC 240
Qy 241 GGGAAATGATTTTGTAGCATCGAAACACAGGAATCAGGTTGAACGCTCCTCAGATGACT 300
Db 241 GGGAAATGATTTTGTAGCATCGAAACACAGGAATCAGGTTGAACGCTCCTCAGATGACT 300
Qy 301 TTCCGACGCTCCAGAGAAATCTTCCGAGAGATCATGCCAAGAGCCAGCAGAGAGAA 360
Db 301 TTCCGACGCTCCAGAGAAATCTTCCGAGAGATCATGCCAAGAGCCAGCAGAGAGAA 360
Qy 361 AATGTTTGAAGGAAGTCCAGAGCATCTGCGCCACAAAATGATGGAAACAGCTGTGC 420
Db 361 AATGTTTGAAGGAAGTCCAGAGCATCTGCGCCACAAAATGATGGAAACAGCTGTGC 420
Qy 421 CCCCAGGAAATCCAGTACCTTGGAGAGATTAACAGACATCTGGACCCAAAGGGGG 480
Db 421 CCCCAGGAAATCCAGTACCTTGGAGAGATTAACAGACATCTGGACCCAAAGGGGG 480
Qy 481 AAACATGCTCGGACCCACAGACTGCTGAGAGAAAAGCAGCTGGTGGTTTATGAAGATC 540
Db 481 AAACATGCTCGGACCCACAGACTGCTGAGAGAAAAGCAGCTGGTGGTTTATGAAGATC 540

QY 181 CTAGGTTTCAAGTCCACCCCTTTCATCGTAGTAAACGGGCTGCAGACTTCCAC 240
 Db 181 CTAGGTTTCAAGTCCACCCCTTTCATCGTAGTAAACGGGCTGCAGACTTCCAC 240
 QY 241 GGGAAATGATTTGGTAAAGTTCGAAACACAGGAATCAGGTTGAAAGTCTCAGATGACT 300
 Db 241 GGGAAATGATTTGGTAAAGTTCGAAACACAGGAATCAGGTTGAAAGTCTCAGATGACT 300
 QY 301 TTGGGACGCTCCAGAGAAATCTTCCGGAAGATCATGCCCAAGAGCCAGCAGAGGAAGAA 360
 Db 301 TTGGGACGCTCCAGAGAAATCTTCCGGAAGATCATGCCCAAGAGCCAGCAGAGGAAGAA 360
 QY 361 AATCGTTGAAGGAAGTCCAGAGGATCTGGGCCCAAAAATGATGGGAACAGCTGTGC 420
 Db 361 AATCGTTGAAGGAAGTCCAGAGGATCTGGGCCCAAAAATGATGGGAACAGCTGTGC 420
 QY 421 CCCCCGGAATCAAGTACCTTCGAGAAGATTACAGACATCTGGACCCCAAAAGGGG 480
 Db 421 CCCCCGGAATCAAGTACCTTCGAGAAGATTACAGACATCTGGACCCCAAAAGGGG 480
 QY 481 AAACATGCTCGACCCACAGACTCGTGGAGAAAGCAGCTGGTGTGTTATGAAGATC 540
 Db 481 AAACATGCTCGACCCACAGACTCGTGGAGAAAGCAGCTGGTGTGTTATGAAGATC 540
 QY 541 AGCGACCTCGAGGAGATGACGAGTAACTCCCTCG 576
 Db 541 AGCGACCTCGAGGAGATGACGAGTAACTCCCTCG 576

BC005325 1250 bp mRNA linear PRI 03-OCT-2003
 Homo sapiens synovial sarcoma, X breakpoint 4, transcript variant
 1, mRNA (CDNA clone MGC:12411 IMAGE:3961688), complete cds.
 BC005325
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)

REFERENCE
 AUTHORS
 Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
 Klausner, R.D., Collins, P.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
 Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
 Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
 Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
 Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
 Abramson, R.D., Mullah, S.J., Bosak, S.A., McEwan, P.J.,
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
 Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
 Fahey, J., Helton, E., Kettner, M., Madan, A., Rodriguez, S.,
 Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y.,
 Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
 Butterfield, Y.S., Krzywinski, M.I., Skalek, U., Smalish, D.E.,
 Scherch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

TITLE
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 1 (bases 1 to 1250)
 2 (bases 1 to 1250)
 Strausberg, R.
 Direct Submission
 Submitted (27-MAR-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>

REMARK

COMMENT

Contact: MGC help desk
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: CLONTECH Laboratories, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILLNL)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www-shgc.stanford.edu>
 Contact: (Dickson, Mark) mcd@paxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
 R. M.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/ILLNL at: <http://image.llnl.gov>
 Series: IMAGE Plate: 16 Row: m Column: 20
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 28559011.

FEATURES

Location/Qualifiers
 1..1250
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:12411 IMAGE:3961688"
 /tissue_type="Bladder, carcinoma"
 /clone_lib="NIH MGC_53"
 /lab_host="DH10B"
 /note="Vector: pDNR-LIB"
 1..1250
 /gene="SSX4"
 /note="synonym: MGC12411"
 /db_xref="LocustID:6759"
 /db_xref="MIM:300326"
 59..625
 /codon_start=1
 /product="synovial sarcoma, X breakpoint 4, isoform a"
 /protein_id="AA05325.1"
 /db_xref="GI:13529095"
 /db_xref="LocustID:6759"
 /translation="MNGDDAFRRPRDDAQISEKLKAFDDIAKYFSKEWEKMSKSE
 KIVYVMKLNVTMKGKLVLPFPMRSKAAADFGNDGNDNRHNRQVPRPQTFG
 SLQRIFFPKPKPAEENGKEVPEASGPQNDGKQLCPGPNPSTLEKINTKSGPKRG
 KXANTHLRERKQLVYVEISPEEDDE"
 125..304
 /note="KXAB; Region: krueppel associated box"
 /db_xref="CDD:smart00349"

misc_feature

ORIGIN

Query March 99.7%; Score 574.4; DB 9; Length 1250;
 Best Local Similarity 99.8%; Pred. No. 8.8e-164;
 Matches 575; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 ATGAACGGAGACGACCCCTTTGCAAGAGACCCAGGATGCTCAATATCAGAGAAG 60
 Db 59 ATGAACGGAGACGACCCCTTTGCAAGAGACCCAGGATGCTCAATATCAGAGAAG 118
 Qy 61 TTACGAAAGCCTTCGATGATATGCGCAATATCTTCTAAGAAAGAGTGGGAAAGATG 120
 Db 119 TTACGAAAGCCTTCGATGATATGCGCAATATCTTCTAAGAAAGAGTGGGAAAGATG 178
 Qy 121 AAATCTCTCGAGAAAAATCGCTATGTGTATATGAAGCTAAACTATCAGGTCTATGACTAA 180
 Db 179 AAATCTCTCGAGAAAAATCGCTATGTGTATATGAAGCTAAACTATCAGGTCTATGACTAA 238
 Qy 181 CTAGGTTTCAAGTCCACCCCTTTCATCGTAGTAAACGGGCTGCAGACTTCCAC 240
 Db 239 CTAGGTTTCAAGTCCACCCCTTTCATCGTAGTAAACGGGCTGCAGACTTCCAC 298
 Qy 241 GGGAAATGATTTGGTAAAGTTCGAAACACAGGAATCAGGTTGAAAGTCTCAGATGACT 300
 Db 299 GGGAAATGATTTGGTAAAGTTCGAAACACAGGAATCAGGTTGAAAGTCTCAGATGACT 358
 Qy 301 TTGGGACGCTCCAGAGAAATCTTCCGGAAGATCATGCCCAAGAGCCAGCAGAGGAAGAA 360

Db 359 TTGGCAGCCTCCAGAGAAATCTTCCGAGATCATGCCCAAGACGAGAGAGAA 418
QY 361 AATGGTTTGAAGAGTGCAGAGGATCTGCCCAAAAATGATGGGAAACAGCTGTGC 420
Db 419 AATGGTTTGAAGAGTGCAGAGGATCTGCCCAAAAATGATGGGAAACAGCTGTGC 478
QY 421 CCCCCGGGAAATCCAAAGTACCTTGGAGAGATTACAGACATCTGGACCCAAAAGGGGG 480
Db 479 CCCCCGGGAAATCCAAAGTACCTTGGAGAGATTACAGACATCTGGACCCAAAAGGGGG 538
QY 481 AAACATGCTTGAGCCACAGACTGCGTGAGAGAAAGCAGCTGGTGGTTTATGAAGAGATC 540
Db 539 AAACATGCTTGAGCCACAGACTGCGTGAGAGAAAGCAGCTGGTGGTTTATGAAGAGATC 598
QY 541 AGCAGCCCTGAGAGAGATGACGAGTAACCTCCCTCG 576
Db 599 AGCAGCCCTGAGAGAGATGACGAGTAACCTCCCTCG 634

RESULT 10
A48452
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source

A48452
Sequence 17 from Patent WO9602641.
A48452
A48452.1 GI:2302236
unidentified
unclassified.
1 (bases 1 to 766)
Cooper, C.S. and Gusterson, B.A.
MATERIALS AND METHODS RELATING TO THE DIAGNOSIS AND PROPHYLACTIC
AND THERAPEUTIC TREATMENT OF SYNOVIAL SARCOMA
Patent: WO 9602641-A 17 01-FEB-1996;
CANCER RES CAMPAIGN TECH (GB)
Other publication AU 2986595 960216.
Location/Qualifiers
1. .766
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"

ORIGIN
Query Match 86.2%; Score 496.6; DB 6; Length 766;
Best Local Similarity 91.5%; Pred. No. 4.6e-140;
Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 1 ATGAACGGAGACGACGCTTTGCAAGGAGACCCAGGATGATGCTCAAAATACAGAGAG 60
Db 92 ATGAACGGAGACGACGCTTTGCAAGGAGACCCAGGATGATGCTCAAAATACAGAGAG 151
QY 61 TTACGAAAGGCTTCGATGATATTGCCAAATACCTTCTAAGAAAGAGTGGGAAAAGATG 120
Db 152 ATCCAAAAGGCTTCGATGATATTGCCAAATACCTTCTAAGAAAGAGTGGGAAAAGATG 211
QY 121 AAATCCTCGGAGAAAATCGTCTATGTATATGAAGCTAAACTATGAGGTCATGACTAAA 180
Db 212 AAAGCCTCGGAGAAAATCTTCTATGTATATGAAGCTAAACTATGAGGTCATGACTAAA 271
QY 181 CTAGGTTTCAAGTCCACCTCCACCTTTCATCGTAGTAAACCGGCTGCAGACTTCCAC 240
Db 272 CTAGGTTTCAAGTCCACCTCCACCTTTCATCGTAGTAAACCGGCTGCAGACTTCCAC 331
QY 241 GGAATGATTTGGTAAAGTCCAGAGGATCATGCCAGAGTCAATCGCTCCTCAGATGACT 300
Db 332 GGAATGATTTGGTAAAGTCCAGAGGATCATGCCAGAGTCAATCGCTCCTCAGATGACT 391
QY 241 GGAATGATTTGGTAAAGTCCAGAGGATCATGCCAGAGTCAATCGCTCCTCAGATGACT 300
Db 332 GGAATGATTTGGTAAAGTCCAGAGGATCATGCCAGAGTCAATCGCTCCTCAGATGACT 391
QY 301 TTCCGAGCCTCCAGAGAAATCTTCCGAGATCATGCCAGAGTCAATCGCTCCTCAGATGACT 360
Db 392 TTCCGAGCCTCCAGAGAAATCTTCCGAGATCATGCCAGAGTCAATCGCTCCTCAGATGACT 451
QY 361 AATGGTTTGAAGAGTGCAGAGGATCTTGGCCCAAAAATGATGGGAAACAGCTGTGC 420
Db 392 TTCCGAGCCTCCAGAGAAATCTTCCGAGATCATGCCAGAGTCAATCGCTCCTCAGATGACT 451
QY 361 AATGGTTTGAAGAGTGCAGAGGATCTTGGCCCAAAAATGATGGGAAACAGCTGTGC 420

Db 452 AATGATTCGGAGGAAGTCCAGAAAGCATCTGGCCCAAAAATGATGGGAAAGAGCTGTGC 511
QY 421 CCCCCGGGAAATCCAAAGTACCTTGGAGAGATTAAACAAGACATCTGGACCCAAAAGGGGG 480
Db 512 CCCCCGGGAAATCCAAAGTACCTTGGAGAGATTAAACAAGACATCTGGACCCAAAAGGGGG 571
QY 481 AAACATGCTTGAGCCACAGACTGCGTGAGAGAAAGCAGCTGGTGGTTTATGAAGAGATC 540
Db 572 GAACATGCTTGAGCCACAGACTGCGTGAGAGAAAGCAGCTGGTGGTTTATGAAGAGATC 631
QY 541 AGCAGCCCTGAGAGAGATGACGAGTAACCTCCCTCG 575
Db 632 AGCAGCCCTGAGAGAGATGACGAGTAACCTCCCTCG 666

RESULT 11
AX821947
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source

AX821947
Sequence 75 from Patent WO03068961.
AX821947
AX821947.1 GI:39725168
Homo sapiens (human)
Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Andrews, P.A., Walsh, J.A. and Gokhale, P.A.
Method to modify differentiation of pluripotential stem cells
Patent: WO 03068961-A 75 21-AUG-2003;
Axordia Limited (GB)
Location/Qualifiers
1. .766
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 86.2%; Score 496.6; DB 6; Length 766;
Best Local Similarity 91.5%; Pred. No. 4.6e-140;
Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 1 ATGAACGGAGACGACGCTTTGCAAGGAGACCCAGGATGATGCTCAAAATACAGAGAG 60
Db 92 ATGAACGGAGACGACGCTTTGCAAGGAGACCCAGGATGATGCTCAAAATACAGAGAG 151
QY 61 TTACGAAAGGCTTCGATGATATTGCCAAATACCTTCTAAGAAAGAGTGGGAAAAGATG 120
Db 152 ATCCAAAAGGCTTCGATGATATTGCCAAATACCTTCTAAGAAAGAGTGGGAAAAGATG 211
QY 121 AAATCCTCGGAGAAAATCGTCTATGTATATGAAGCTAAACTATGAGGTCATGACTAAA 180
Db 212 AAAGCCTCGGAGAAAATCTTCTATGTATATGAAGCTAAACTATGAGGTCATGACTAAA 271
QY 181 CTAGGTTTCAAGTCCACCTCCACCTTTCATCGTAGTAAACCGGCTGCAGACTTCCAC 240
Db 272 CTAGGTTTCAAGTCCACCTCCACCTTTCATCGTAGTAAACCGGCTGCAGACTTCCAC 331
QY 241 GGAATGATTTGGTAAAGTCCAGAGGATCATGCCAGAGTCAATCGCTCCTCAGATGACT 300
Db 332 GGAATGATTTGGTAAAGTCCAGAGGATCATGCCAGAGTCAATCGCTCCTCAGATGACT 391
QY 301 TTCCGAGCCTCCAGAGAAATCTTCCGAGATCATGCCAGAGTCAATCGCTCCTCAGATGACT 360
Db 392 TTCCGAGCCTCCAGAGAAATCTTCCGAGATCATGCCAGAGTCAATCGCTCCTCAGATGACT 451
QY 361 AATGGTTTGAAGAGTGCAGAGGATCTTGGCCCAAAAATGATGGGAAACAGCTGTGC 420
Db 452 AATGATTCGGAGGAAGTCCAGAGGATCATGCCAGAGTCAATCGCTCCTCAGATGACT 511
QY 421 CCCCCGGGAAATCCAAAGTACCTTGGAGAGATTAAACAAGACATCTGGACCCAAAAGGGGG 480
Db 512 CCCCCGGGAAATCCAAAGTACCTTGGAGAGATTAAACAAGACATCTGGACCCAAAAGGGGG 571

Db 339 GGGAAATGTTGGATAATGACCCCTAACCGTGGGAATCAGTTTGAACGTCCTCAGATGACT 398
QY 301 TTCGCGAGCTCCAGAGATCTTCGGAGATCATGCCAGATCATGCCAGAGCCAGCAGAGAGAA 360
Db 399 TTCGCGAGCTCCAGGAATCTCCCGAAGATCATGCCAGATCATGCCAGAGCCAGCAGAGAGAA 458
QY 361 AATGGTTTGAAGAAAGTGCAGAGGATCTGCGCCACAAAATGATGGGAAACAGCTGTGC 420
Db 459 AATGATTCGGAGGAAGTGCAGAGATCTGCGCCACAAAATGATGGGAAACAGCTGTGC 518
QY 421 CCCCCGGGAATCCAGTACCTTGGAGAGATTAACAGACATCTGGACCCAAAAGGGG 480
Db 519 CCCCCGGGAATCCAGTACCTTGGAGAGATTAACAGACATCTGGACCCAAAAGGGG 578
QY 481 AAACATGCTCGGACCCACAGACTGCGTGAGAGAAAGCAGCTGGTGTATTAAGAGATC 540
Db 579 GAACATGCTCGGACCCACAGACTGCGTGAGAGAAAGCAGCTGGTGTATTAAGAGATC 638
QY 541 ACGGACCTCGAGAGATCAGAGTAACCTCCCTC 575
Db 639 ACGGACCTCGAGAGATCAGAGTAACCTCCCTC 673

RESULT 14
AR060380
LOCUS AR060380 931 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 2 from patent US 5840568.
ACCESSION AR060380
VERSION AR060380.1 GI:5986830
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 931)
AUTHORS Firenzuolu, M.
TITLE Hodgkin's disease associated molecules and uses thereof
JOURNAL Patent: US 5840568-A 2 24-NOV-1998;
FEATURES
Location/Qualifiers
1..931
source /organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 86.2%; Score 496.6; DB 6; Length 931;
Best Local Similarity 91.5%; Pred. No. 4.7e-140;
Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 1 ATGAACGAGACGACGCTTTCAGAGAGACCCAGGATGATGCTCAATATCAGAGAAG 60
Db 99 ATGAACGAGACGACGCTTTCAGAGAGACCCAGGATGATGCTCAATATCAGAGAAG 158
QY 61 TTACGAAAGGCTTCGATGATATTCGCAATATCTCTAAGAAAGAGTGGGAAAGATG 120
Db 159 ATCCAAAGGCTTCGATGATATTCGCAATATCTCTAAGAAAGAGTGGGAAAGATG 218
QY 121 AAATCTCGAGAAATCGTCTATGTATATGAAGTAACTATGAGTCACTATAA 180
Db 219 AAAGCTTCGAGAAATCTTCTATGTATATGAAGTAACTATGAGTCACTATAA 278
QY 181 CTAGGTTTCAAGTCACTCCCTCCACCTTCATCGTAGTAAACGGCTGCAGACTTCCAC 240
Db 279 CTAGGTTTCAAGTCACTCCCTCCACCTTCATGTATTAACGGCCGAGACTTCCAG 338
QY 241 GGAATGATTTGGTAAACGATCGAAACCCAGGATCAGGTTGAACGCTCCTCAGATGACT 300
Db 339 GGAATGATTTGGTAAACGATCGAAACCCAGGATCAGGTTGAACGCTCCTCAGATGACT 398
QY 301 TTGCGAGCTCCAGAGATCTTCCGAGATCATGCCAGAGATCATGCCAGAGAGAA 360
Db 399 TTGCGAGCTCCAGGAATCTCCCGAAGATCATGCCAGAGATCATGCCAGAGAGAA 458
QY 361 AATGGTTTGAAGAAAGTGCAGAGGATCTGGCCCAAAAATGATGGGAAACAGCTGTGC 420
Db 459 AATGATTCGGAGGAAGTGCAGAGATCTGGCCCAAAAATGATGGGAAACAGCTGTGC 518
QY 421 CCCCCGGGAATCCAGTACCTTGGAGAGATTAACAGACATCTGGACCCAAAAGGGG 480
Db 519 CCCCCGGGAATCCAGTACCTTGGAGAGATTAACAGACATCTGGACCCAAAAGGGG 578
QY 481 AAACATGCTCGGACCCACAGACTGCGTGAGAGAAAGCAGCTGGTGTATTAAGAGATC 540

Db 459 AATGATTCGGAGGAAGTGCAGAGATCTGGCCCAAAAATGATGGGAAACAGCTGTGC 518
QY 421 CCCCCGGGAATCCAGTACCTTGGAGAGATTAACAGACATCTGGACCCAAAAGGGG 480
Db 519 CCCCCGGGAATCCAGTACCTTGGAGAGATTAACAGACATCTGGACCCAAAAGGGG 578
QY 481 AAACATGCTCGGACCCACAGACTGCGTGAGAGAAAGCAGCTGGTGTATTAAGAGATC 540
Db 579 GAACATGCTCGGACCCACAGACTGCGTGAGAGAAAGCAGCTGGTGTATTAAGAGATC 638
QY 541 ACGGACCTCGAGAGATCAGAGTAACCTCCCTC 575
Db 639 ACGGACCTCGAGAGATCAGAGTAACCTCCCTC 673

RESULT 15
AR117874
LOCUS AR117874 931 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 2 from patent US 6140464.
ACCESSION AR117874
VERSION AR117874.1 GI:14098780
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 931)
AUTHORS Firenzuolu, M. and Rammensee, H.-G.
TITLE Nonapeptides that bind a HLA-A2.1 molecule
JOURNAL Patent: US 6140464-A 2 31-OCT-2000;
FEATURES
Location/Qualifiers
1..931
source /organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 86.2%; Score 496.6; DB 6; Length 931;
Best Local Similarity 91.5%; Pred. No. 4.7e-140;
Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 1 ATGAACGAGACGACGCTTTCAGAGAGACCCAGGATGATGCTCAATATCAGAGAAG 60
Db 99 ATGAACGAGACGACGCTTTCAGAGAGACCCAGGATGATGCTCAATATCAGAGAAG 158
QY 61 TTACGAAAGGCTTCGATGATATTCGCAATATCTCTAAGAAAGAGTGGGAAAGATG 120
Db 159 ATCCAAAGGCTTCGATGATATTCGCAATATCTCTAAGAAAGAGTGGGAAAGATG 218
QY 121 AAATCTCGAGAAATCGTCTATGTATATGAAGTAACTATGAGTCACTATAA 180
Db 219 AAAGCTTCGAGAAATCTTCTATGTATATGAAGTAACTATGAGTCACTATAA 278
QY 181 CTAGGTTTCAAGTCACTCCCTCCACCTTCATCGTAGTAAACGGCTGCAGACTTCCAC 240
Db 279 CTAGGTTTCAAGTCACTCCCTCCACCTTCATGTATTAACGGCCGAGACTTCCAG 338
QY 241 GGAATGATTTGGTAAACGATCGAAACCCAGGATCAGGTTGAACGCTCCTCAGATGACT 300
Db 339 GGAATGATTTGGTAAACGATCGAAACCCAGGATCAGGTTGAACGCTCCTCAGATGACT 398
QY 301 TTGCGAGCTCCAGAGATCTTCCGAGATCATGCCAGAGATCATGCCAGAGAGAA 360
Db 399 TTGCGAGCTCCAGGAATCTCCCGAAGATCATGCCAGAGATCATGCCAGAGAGAA 458
QY 361 AATGGTTTGAAGAAAGTGCAGAGGATCTGGCCCAAAAATGATGGGAAACAGCTGTGC 420
Db 459 AATGATTCGGAGGAAGTGCAGAGATCTGGCCCAAAAATGATGGGAAACAGCTGTGC 518
QY 421 CCCCCGGGAATCCAGTACCTTGGAGAGATTAACAGACATCTGGACCCAAAAGGGG 480
Db 519 CCCCCGGGAATCCAGTACCTTGGAGAGATTAACAGACATCTGGACCCAAAAGGGG 578
QY 481 AAACATGCTCGGACCCACAGACTGCGTGAGAGAAAGCAGCTGGTGTATTAAGAGATC 540

Db 579 GAACATGCCCTGGACCCACAGACTGCGTGAGAGAAAACAGCTGGTGATTATGAAGAGATC 638

QY 541 AGCGACCCCTGAGGAAGATGACGAGTAACTCCCTC 575

Db 639 AGCGACCCCTGAGGAAGATGACGAGTAACTCCCTC 673

Search completed: March 31, 2004, 16:39:15
Job time : 3494 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 31, 2004, 13:52:29 ; Search time 437 seconds
(without alignment)
5599.456 Million cell updates/sec

Title: US-09-975-856-1

Perfect score: 576

Sequence: 1 ATGACGAGGACGACGCTT.....ATGACGAGTAACTCCCTCG 576

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_29Jan04.*

- 1: Geneseqn1980s.*
- 2: Geneseqn1990s.*
- 3: Geneseqn2000s.*
- 4: Geneseqn2001as.*
- 5: Geneseqn2001bs.*
- 6: Geneseqn2002as.*
- 7: Geneseqn2003as.*
- 8: Geneseqn2003bs.*
- 9: Geneseqn2003cs.*
- 10: Geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	576	100.0	576	6	ABK84472 Human cDN
2	576	100.0	576	7	ABQ83858 Human SSX
3	576	100.0	1250	9	ADC09598
4	574.4	99.7	576	2	AAV70060 Human SSX
5	496.6	86.2	766	2	AAT11780 Human X-c
6	496.6	86.2	766	6	ABS73282 DNA encod
7	496.6	86.2	766	7	AAD54033 Human col
8	496.6	86.2	766	7	ABQ83844 Human SSX
9	496.6	86.2	766	9	ADC09570 SSX-2 cDN
10	496.6	86.2	766	9	ADD25523 Binding d
11	496.6	86.2	931	2	AAT47748 Melanoma
12	496.6	86.2	931	2	AAV04267 Melanoma
13	496.6	86.2	1309	6	ABL64122 Breast ca
14	496.6	86.2	1309	6	ABL63730 Breast ca
15	487.2	84.6	921	5	AAS80402 DNA encod
16	487.2	84.6	921	5	AAS80399 DNA encod
17	480.6	83.4	576	2	AAV70061 Human SSX
18	480	83.3	766	2	AAT11779 Human X-c
19	221.4	38.4	711	2	AAT11781 Human SVT
20	209.6	36.4	585	2	AAT11782 Human SVT
21	192.6	33.4	830	5	AAS80401 DNA encod
22	191.4	33.2	822	5	AAS80404 DNA encod
23	117.4	20.4	165	6	ABS73281 DNA encod

24	74.2	12.9	1110	5	AAS91978	Aas91978 DNA encod
25	42.2	7.3	2000	7	ADA71938	Ada71938 Rice gene
26	41.4	7.2	1981	9	ADC32171	Adc32171 Human nov
27	40.8	7.1	80	6	ABN33980	Abn33980 Human spl
28	40.8	7.1	8059	3	AAAB1747	Aaas1747 N. mening
29	40.8	7.1	110000	3	AAAB1490-07	Continuation 18 of
30	40.8	7.1	349980	3	AAF21608	Aaf21608 Neisseria
31	39.8	6.9	571	6	ABK71614	Abk71614 Human dit
32	39.8	6.9	3591	8	ADA10966	Ada10966 Human cDN
33	39.8	6.9	3591	8	ACH04006	Ach04006 Human cDN
34	39.4	6.8	1097	4	AAF27660	Aaf27660 DNA encod
35	39.4	6.8	1097	4	ACH04007	Ach04007 Human cDN
36	37.8	6.6	2736	4	AAI57976	Aai57976 Human pol
37	37.8	6.6	2872	4	AAH98654	Aah98654 Human EST
38	37.8	6.6	2872	4	AAI59762	Aai59762 Human pol
39	37.8	6.6	5532	9	ADB47398	Adb47398 Human cDN
40	37.4	6.5	730	4	AAC91340	Aac91340 Human pol
41	37	6.4	653	6	ABQ99153	Abq99153 Human ORF
42	36	6.2	2369	6	ABA01156	Ab01156 Human zin
43	36	6.2	2798	9	ADC30277	Adc30277 Human nov
44	36	6.2	349980	5	AAH41224	Aah41224 Pyrococcu
45	35.8	6.2	765	4	AAK91889	Aak91889 Human cDN

ALIGNMENTS

RESULT 1
ABK84472
ID ABK84472 standard; cDNA; 576 BP.
XX
AC ABK84472;
XX
DT 14-AUG-2002 (first entry)
XX
DE Human cDNA differentially expressed in granulocytic cells #1043.
XX
KW Human; ss; granulocytic cell; DNA chip; bacterial infection;
KW viral infection; parasitic infection; protozoal infection;
KW fungal infection; sterile inflammatory disease; psoriasis;
KW rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
KW cardiac reperfusion injury; renal reperfusion injury; AKDS;
KW adult respiratory distress syndrome; inflammatory bowel disease;
KW Crohn's disease; ulcerative colitis; periodontal disease;
KW granulocyte activation; chronic inflammation; allergy.
XX
OS Homo sapiens.
XX
PN WO200228999-A2.
XX
PD 11-APR-2002.
PF 03-OCT-2001; 2001WO-US030821.
XX
PR 03-OCT-2000; 2000US-0237189P.
XX
(GENE-) GENE LOGIC INC.
XX
PI Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;
XX
DR WPI; 2002-435328/46.
XX
PT Detecting granulocyte activation by detecting differential expression of
genes associated with granulocyte activation, which serves as diagnostic
markers that is useful for monitoring disease states and drug toxicity.
XX
PS Claim 1; SEQ ID NO 1043; 114pp; English.
XX
CC The invention relates to detecting (M1) granulocyte (GC) activation
(GCA), by detecting the level of expression of gene(s) (Gs) identified by
DNA chip analysis as given in the specification, and comparing the
expression level to an expression level in an unactivated GC, where
differential expression of Gs is indicative of GCA. Also included are

CC modulating (M2) GA by contacting GC with an agent that alters the
CC expression of at least one gene in Gs; (2) screening (M3) for an agent
CC capable of modulating GCA or an inflammation (especially chronic) in a
CC tissue, an allergic response in a subject, exposure of a subject to a
CC pathogen or sterile inflammatory disease using the gene expression
CC profile; (3) detecting (M4) an inflammation (especially chronic) in a
CC tissue, an allergic response in a subject, exposure of a subject to a
CC pathogen or sterile inflammatory disease, by detecting the level of
CC expression in a sample of the tissue of gene(s) from Gs, where the level
CC of expression of the gene is indicative of inflammation; (4) treating
CC (M5) an inflammation (especially chronic) or in a tissue, an allergic
CC response in a subject, exposure of a subject to a pathogen or sterile
CC inflammatory disease, by contacting a tissue having inflammation with an
CC agent that modulates the expression of gene(s) from Gs in the tissue. M1
CC is useful for detecting GCA; M2 is useful for modulating GCA; M3 is useful
CC for screening an agent capable of modulating GCA preferably in an
CC inflammation in a tissue; M4 is useful for detecting an inflammation
CC (especially chronic) in a tissue, an allergic response in a subject,
CC exposure of a subject to a pathogen or sterile inflammatory disease (e.g.
CC psoriasis, rheumatoid arthritis, glomerulonephritis, asthma, thrombosis,
CC cardiac reperfusion injury, renal reperfusion injury, ARDS, adult
CC respiratory distress syndrome, inflammatory bowel disease, Crohn's
CC disease, ulcerative colitis, periodontal disease, also bacterial
CC infection, viral infection, parasitic infection, protozoal infection,
CC fungal infection and M5 is useful for treating one of the above
CC conditions. The present sequence represents a gene differentially
CC expressed in granulocytes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX

Sequence 576 BP; 187 A; 127 C; 150 G; 112 T; 0 U; 0 Other;
Query Match 100.0%; Score 576; DB 6; Length 576;
Best Local Similarity 100.0%; Pred. No. 9.3e-172;
Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACGAGACGAGCGCTTTCGACGAGACCCAGGATGATGCTCAAAATATCAGAGAAG 60
DB 1 ATGACGAGACGAGCGCTTTCGACGAGACCCAGGATGATGCTCAAAATATCAGAGAAG 60
QY 61 TTACGAAAGGCGCTTCGATGATATTGCCAAATPACTTCTTAAGAAAGAGTGGAAAGATG 120
DB 61 TTACGAAAGGCGCTTCGATGATATTGCCAAATPACTTCTTAAGAAAGAGTGGAAAGATG 120
QY 121 AAATCTCGGAGAAATCGTCTATGTGTATGATGAAGCTAACTATGAGGTCATGACTAAA 180
DB 121 AAATCTCGGAGAAATCGTCTATGTGTATGATGAAGCTAACTATGAGGTCATGACTAAA 180
QY 181 CTAGGTTTCAAGGTCAAGCTCCACCTTTCATGCTGATGATGATGATGATGATGATGATG 240
DB 181 CTAGGTTTCAAGGTCAAGCTCCACCTTTCATGCTGATGATGATGATGATGATGATGATG 240
QY 241 GGAATGATTTTGGTAAGATCGAATGGAACACAGGATGAGTTGACGTCTCAGATGACT 300
DB 241 GGAATGATTTTGGTAAGATCGAATGGAACACAGGATGAGTTGACGTCTCAGATGACT 300
QY 301 TTCGCGAGCTCCAGAGATCTTCCGAGAGATCATGCCCAAGAGCCAGCAGAGAGAGAA 360
DB 301 TTCGCGAGCTCCAGAGATCTTCCGAGAGATCATGCCCAAGAGCCAGCAGAGAGAGAA 360
QY 361 AATGTTTGAAGAGATGCGAGAGATCTGSCCCACAAATATGATGGGAACAGCTGTGC 420
DB 361 AATGTTTGAAGAGATGCGAGAGATCTGSCCCACAAATATGATGGGAACAGCTGTGC 420
QY 421 CCCCGGGAATCCAGTACCTTGGAGAGATTTAACAAGACATCTGGACCCCAAAAGGGGG 480
DB 421 CCCCGGGAATCCAGTACCTTGGAGAGATTTAACAAGACATCTGGACCCCAAAAGGGGG 480
QY 481 AAACATGCTCGACCCACAGATGCTGGTGAAGAGAAAGCAGCTGGTGGTTTATGAAGATC 540
DB 481 AAACATGCTCGACCCACAGATGCTGGTGAAGAGAAAGCAGCTGGTGGTTTATGAAGATC 540

QY 541 AGGACCGCTCAGGAAGATGACGAGTAACCTCCCTCG 576
DB 541 AGGACCGCTCAGGAAGATGACGAGTAACCTCCCTCG 576
RESULT 2
ABQ83858
ID ABQ83858 standard; cDNA; 576 BP.
AC ABQ83858;
XX
XX 03-FEB-2003 (first entry)
XX Human SSX-4 encoding cDNA SEQ ID NO:599.
XX Human; epitope; vaccine; immunotherapeutic; cytostatic; immunogenicity;
KW T cell; gene; ss.
XX Homo sapiens.
XX WO200281646-A2.
XX 17-OCT-2002.
XX 04-APR-2002; 2002WO-US011101.
XX 06-APR-2001; 2001US-0282211P.
XX 07-NOV-2001; 2001US-0337017P.
XX 07-MAR-2002; 2002US-0363210P.
XX (CTLI-) CTL IMMUNOTHERAPIES CORP.
XX Simard JLL, Diamond DC, Liu L, Xie Z;
XX WPI; 2003-067518/06.
XX P-PSDB; ABP74710.
XX Novel epitopes useful as vaccines, comprises peptides or nucleic acid
XX encoding the peptides, that are useful epitopes of target-associated
XX antigens.

Claim 1; Page 185; 352pp; English.
The present invention describes an isolated epitope (I) and an epitope
cluster. Also described is a vaccine or immunotherapeutic composition
(VC) comprising (I). (I) has cytostatic activity. VC is useful for
treating an animal, by administering to an animal the vaccine or
immunotherapeutic composition. VC is also useful for evaluating
immunogenicity of a vaccine or immunotherapeutic composition, by
administering VC to an HLA-transgenic animal and evaluating
immunogenicity based on a characteristic of the animal, or by in vitro
primary stimulation of a T cell and evaluating immunogenicity. (I) is
useful for determining specific T cell frequency, by contacting T cells
with a MHC-peptide complex, and further comprises ELISPOT analysis,
limiting dilution analysis, flow cytometry, in situ hybridisation and/or
polymerase chain reaction (PCR). ABQ83858 and ABP74128 to
ASP74713 represent sequences used in the exemplification of the present
invention

Sequence 576 BP; 187 A; 127 C; 150 G; 112 T; 0 U; 0 Other;
Query Match 100.0%; Score 576; DB 7; Length 576;
Best Local Similarity 100.0%; Pred. No. 9.3e-172;
Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGACGAGACGAGCGCTTTCGACGAGACCCAGGATGATGCTCAAAATATCAGAGAAG 60
DB 1 ATGACGAGACGAGCGCTTTCGACGAGACCCAGGATGATGCTCAAAATATCAGAGAAG 60
QY 61 TTACGAAAGGCGCTTCGATGATATTGCCAAATPACTTCTTAAGAAAGAGTGGAAAGATG 120
DB 61 TTACGAAAGGCGCTTCGATGATATTGCCAAATPACTTCTTAAGAAAGAGTGGAAAGATG 120

QY 121 AAATCTCGGAGAAATCTCTATGTATATGAAGCTAAACTATGAGTTCATGACTAAA 180
DB 121 AAATCTCGGAGAAATCTCTATGTATATGAAGCTAAACTATGAGTTCATGACTAAA 180
QY 181 CTAGGTTTCAAGTTCACCTCCACCTTTTCATGCGTAGTAAACGGGCTGCAGACTTCCAC 240
DB 181 CTAGGTTTCAAGTTCACCTCCACCTTTTCATGCGTAGTAAACGGGCTGCAGACTTCCAC 240
QY 241 GGAATGATTTGGTAAAGTGAACATGAACACACAGGAATCAGGTTGAACGTCCTCAGATGACT 300
DB 241 GGAATGATTTGGTAAAGTGAACATGAACACACAGGAATCAGGTTGAACGTCCTCAGATGACT 300
QY 301 TTCCGAGCTCCAGAGATCTTCCCGAAGATCATGCCCAAGAACCCAGCAGAGAGAA 360
DB 301 TTCCGAGCTCCAGAGATCTTCCCGAAGATCATGCCCAAGAACCCAGCAGAGAGAA 360
QY 361 AATGTTTGAAGAAATGCGCAGAGGATCTGCGCCCAAAATATGAGGAAACAGCTGTGC 420
DB 361 AATGTTTGAAGAAATGCGCAGAGGATCTGCGCCCAAAATATGAGGAAACAGCTGTGC 420
QY 421 CCCCCGGAAATCCAAAGTACCTTGGAGAGATTAAACAGACATCTGGACCCAAAGGGGG 480
DB 421 CCCCCGGAAATCCAAAGTACCTTGGAGAGATTAAACAGACATCTGGACCCAAAGGGGG 480
QY 481 AAACATGCTTGGAGGACCCACAGACTGCGTGAGAGAAACAGCTGTGTTTATGAAGATC 540
DB 481 AAACATGCTTGGAGGACCCACAGACTGCGTGAGAGAAACAGCTGTGTTTATGAAGATC 540
QY 541 AGCGACCTGAGGAGATGACGAGTAACCTCCCTCG 576
DB 541 AGCGACCTGAGGAGATGACGAGTAACCTCCCTCG 576

RESULT 3

ADC09598
ID ADC09598 standard; cDNA; 1250 BP.
XX AC ADC09598;
XX AC
DT 18-DEC-2003 (first entry)
XX SSX-4 cDNA #SEQ ID 599.
XX Epitope; immunological; vaccine;
KW major histocompatibility complex class I; MHC class I; cancer;
KW immunisation; ss.
XX Unidentified.
XX WO2003008537-A2.
XX 30-JAN-2003.
XX 29-MAR-2002; 2002WO-US010189.
XX 06-APR-2001; 2001US-0282211P.
XX 07-NOV-2001; 2001US-0337017P.
XX 07-MAR-2002; 2002US-0363210P.
XX (CTL1-) CTL IMMUNOTHERAPIES CORP.
XX Simard JLL, Diamond DC, Liu L, Xie Z;
XX WPI; 2003-248010/24.

Epitope having high affinity for major histocompatibility complex class I
PT useful for treating an animal, evaluating immunogenicity of a vaccine or
PT therapeutic composition and for diagnosing a disease.
XX
PS Claim 1; SEQ ID NO 599; 239pp; English.
XX The invention relates to an isolated epitope polypeptide that has high
CC affinity for major histocompatibility complex (MHC) class I, and an
CC

CC epitope cluster comprising the polypeptide. Also disclosed is a vaccine
CC or immunotherapeutic composition containing an epitope of the invention.
CC Compositions of the invention may be used in the treatment of cancer. The
CC method can be combined with a radiation therapy, chemotherapy,
CC biochemotherapy or surgery. The composition is also useful for evaluating
CC immunogenicity of a vaccine or immunotherapeutic compound. Multimeric MHC
CC-peptide complexes of the invention are useful for determining specific T
CC cell frequency. This method is useful for evaluating immunological
CC response, by performing the method prior to and subsequent to an
CC immunisation step. Compositions of the invention are useful for
CC diagnosing a disease. The current sequence represents an epitope of the
CC invention with high affinity for MHC class I.

XX SQ Sequence 1250 BP; 373 A; 302 C; 278 G; 297 T; 0 U; 0 Other;

Query Match 100.0%; Score 576; DB 9; Length 1250;
Best Local Similarity 100.0%; Pred. No. 1.4e-171;
Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAACGGAGACGACGCTTTGCAAGGAGACCCAGGATGATGCTCAATATCAGAGAAG 60
DB 59 ATGAACGGAGACGACGCTTTGCAAGGAGACCCAGGATGATGCTCAATATCAGAGAAG 118
QY 61 TTACGAAGGCTTCGATGATATTGCCAATACTTCTTAAGAAAGAGTGGAAAGATG 120
DB 119 TTACGAAGGCTTCGATGATATTGCCAATACTTCTTAAGAAAGAGTGGAAAGATG 178
QY 121 AAATCTCGGAGAAATCGTCTATGTGTATATGAAGCTAAACTATGAGTTCATGACTAAA 180
DB 179 AAATCTCGGAGAAATCGTCTATGTGTATATGAAGCTAAACTATGAGTTCATGACTAAA 238
QY 181 CTAGGTTTCAAGGTCACCTCCACCTTTCATGCGTAGTAAACGGGCTGCAGACTTCCAC 240
DB 239 CTAGGTTTCAAGGTCACCTCCACCTTTCATGCGTAGTAAACGGGCTGCAGACTTCCAC 298
QY 241 GGAATGATTTTGGTAAAGTCCAGAGGATCAGGTTGAACGTCCTCAGATGACT 300
DB 299 GGAATGATTTTGGTAAAGTCCAGAGGATCAGGTTGAACGTCCTCAGATGACT 358
QY 301 TTCCGACCTCCAGAGATCTTCCCGAAGATCATGCCAGAGAGCCAGCAGAGAGAGAA 360
DB 359 TTCCGACCTCCAGAGATCTTCCCGAAGATCATGCCAGAGAGCCAGCAGAGAGAGAA 418
QY 361 AATGTTTGAAGGAAAGTCCAGAGGATCTGCGCCCAAAAATGATGGGAAACAGCTGTGC 420
DB 419 AATGTTTGAAGGAAAGTCCAGAGGATCTGCGCCCAAAAATGATGGGAAACAGCTGTGC 478
QY 421 CCCCCGGAAATCCAAAGTACCTTGGAGAGATTAAACAGACATCTGGACCCAAAGGGGG 480
DB 479 CCCCCGGAAATCCAAAGTACCTTGGAGAGATTAAACAGACATCTGGACCCAAAGGGGG 538
QY 481 AAACATGCTGAGACCCACAGACTGCGTGAGAGAAAGCAGCTGTTTATGAAGAGATC 540
DB 539 AAACATGCTGAGACCCACAGACTGCGTGAGAGAAAGCAGCTGTTTATGAAGAGATC 598
QY 541 AGCGACCTGAGGAGATGACGAGTAACCTCCCTCG 576
DB 599 AGCGACCTGAGGAGATGACGAGTAACCTCCCTCG 634

RESULT 4

AAV70060
ID AAV70060 standard; cDNA; 576 BP.
XX AC AAV70060;
XX AC
DT 01-MAR-1999 (first entry)
XX Human SSX4 cDNA.
XX DE
XX SSX4; SSX gene; human; tumour associated antigen; cancer; melanoma; ss.
XX Homo sapiens.
OS

XX WO9850528-A1.
 XX 12-NOV-1998.
 XX 25-FEB-1998; 98WO-US003661.
 XX 05-MAY-1997; 97US-00851138.
 XX (LUDM-) LUDWIG INST CANCER RES.
 XX Gure AO, Tureci O, Sahin U, Tsang S, Scanlan MJ, Knuth A;
 PI Pfeundschnuh M, Old LJ, Chen Y;
 XX WPI; 1998-610379/51.
 XX New SSX gene family members - useful for assaying for cancer cells.
 PS Claim 5; Page 12; 19pp; English.
 XX This is the nucleotide sequence of a newly isolated human SSX4 cDNA
 CC clone. The cDNA was isolated from a human testicular cDNA by PCR
 CC amplification using primers (see AAV70062-63) based on the known SSX2
 CC sequence. A SSX5 clone (see AAV70061) was also obtained. 2 forms of SSX4
 CC were identified. One of these lacked nucleotides 331-456 but was
 CC otherwise identical to the present SSX4 sequence and is described as an
 CC alternatively spliced form. SSX4 shares 89.4% homology to SSX1 on the
 CC nucleotide level and 79.3% at the amino acid level. The inventional
 CC additionally provides expression vectors, transformed or transfected
 CC cells that can be used to produce SSX proteins, and primers (see AAV70062
 CC -73) useful for determining expression of an SSX gene in a sample. The
 CC new SSX genes can be used to assay for cancers such as melanoma
 XX
 SQ Sequence 576 BP; 188 A; 127 C; 149 G; 112 T; 0 U; 0 Other;
 Query Match 99.7%; Score 574.4; DB 2; Length 576;
 Best Local Similarity 99.8%; Pred. No. 3e-171;
 Matches 575; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 ATGACCGAGACGCGCTTGCAGAGAGACCCAGGAGTATCTCTAAGAAAGAGTGGGAAAGATG 120
 DB 1 ATGACCGAGACGCGCTTGCAGAGAGACCCAGGAGTATCTCTAAGAAAGAGTGGGAAAGATG 120
 QY 61 TTACGAAAGGCTTCGATGATATTCGCAATATCTCTAAGAAAGAGTGGGAAAGATG 120
 DB 61 TTACGAAAGGCTTCGATGATATTCGCAATATCTCTAAGAAAGAGTGGGAAAGATG 120
 QY 121 AAATCCTCGAGAAATCGTCTATGTATATGAGCTAACTAATGAGTCACTGACTAA 180
 DB 121 AAATCCTCGAGAAATCGTCTATGTATATGAGCTAACTAATGAGTCACTGACTAA 180
 QY 181 CTAGGTTCAAGGTCACCTCCACCTTCATGCTAGTAAACGGGCTGCAGACTTCCAC 240
 DB 181 CTAGGTTCAAGGTCACCTCCACCTTCATGCTAGTAAACGGGCTGCAGACTTCCAC 240
 QY 241 GGGAAATGTTTGTAAAGTACGATCGAAACACAGGAATCAGGTTGAACTCTCTCAGATGACT 300
 DB 241 GGGAAATGTTTGTAAAGTACGATCGAAACACAGGAATCAGGTTGAACTCTCTCAGATGACT 300
 QY 301 TTCGCGACCTCCAGAGAACTCTCCGAGGATCATGCCAAGAGTCCAGAGCCAGAGGAGAA 360
 DB 301 TTCGCGACCTCCAGAGAACTCTCCGAGGATCATGCCAAGAGTCCAGAGCCAGAGGAGAA 360
 QY 361 AATGTTTGAAGGAGTCCAGAGGATCTGCGCCACAAATATGATGGGAAACAGCTGTGC 420
 DB 361 AATGTTTGAAGGAGTCCAGAGGATCTGCGCCACAAATATGATGGGAAACAGCTGTGC 420
 QY 421 CCCCAGGAATCCAGTACCTTGGAGAGATTACAGACATCTGCAGCCCAAGGGGG 480
 DB 421 CCCCAGGAATCCAGTACCTTGGAGAGATTACAGACATCTGCAGCCCAAGGGGG 480
 QY 481 AAACATGCTCGACCCACAGACTCGGTGAGAGAAAGCAGCTGGTGGTTTATGAAGATC 540

DB 481 AAACATGCTCGACCCACAGACTCGGTGAGAGAAAGCAGCTGGTGGTTTATGAAGATC 540
 QY 541 AGCGACCTCGAGAGATGACGAGTAACCTCCCTCG 576
 DB 541 AGCGACCTCGAGAGATGACGAGTAACCTCCCTCG 576
 RESULT 5
 AAT11780
 ID AAT11780 standard; cDNA; 766 BP.
 XX AAT11780;
 AC AAT11780;
 DT 16-APR-1996 (first entry)
 XX
 DE Human X-chromosome SSX2 cDNA.
 XX Human; SSX2 gene; synovial sarcoma; X-chromosome; breakpoint-2;
 KW ornithine-delta-aminotransferase; OATL2 gene; Xp11.2; SSX1 gene;
 KW translocation; chromosome-18; 18q11.2; SYT gene; gene fusion; SYT-SSX2;
 KW fusion protein; primer; PCR; polymerase chain reaction; Smal; LpsI;
 KW probe; antibody; monoclonal antibody; humanised antibody; hybridisation;
 KW antisense; antitumour; recombinant vaccine; vaccinia virus; vector;
 KW cancer; diagnosis; therapy; ss.
 XX Homo sapiens.
 OS
 XX
 FH Key Location/Qualifiers
 FT misc_difference 230..232
 FT /tag= a
 FT /codon= seq:TTC, aa:Ser
 FT misc_difference 251..253
 FT /tag= b
 FT /codon= seq:AAG, aa:Asn
 FT misc_difference 257..259
 FT /tag= c
 FT /codon= seq:GAG, aa:Lys
 FT misc_difference 377..379
 FT /tag= d
 FT /codon= seq:CGT, aa:Ileu
 FT misc_recomb 421..422
 FT /tag= e
 FT /note= "Breakpoint for SYT translocation"
 FT misc_feature 422..766
 FT /tag= f
 FT /note= "3'-Region present in SYT-SSX2 fusion"
 FT primer_binding 747..766
 FT /tag= g
 FT /note= "Binds primer AAT11784"
 XX WO9602641-A2.
 PN
 XX
 XX 01-FEB-1996.
 XX
 XX 19-JUL-1995; 95WO-GB001704.
 XX
 XX 19-JUL-1994; 94GB-00014580.
 PR
 XX
 PA (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.
 XX
 XX Cooper CS, Gusterson BA;
 PI
 XX WPI; 1996-105904/11.
 DR P-PSDB; AAR90677.
 DR
 XX Diagnosing synovial sarcoma by detection of e.g. SYT and SSX1 sequences,
 PT based on X18 trans-location - by amplification of trans-location
 PT sequences or detection of the encoded protein.
 PT
 XX Claim 6; Fig 7; 55pp; English.
 XX
 CC The sequence represents the SSX2 gene (synovial sarcoma X-chromosome
 CC breakpoint-2). The gene is at the location of a breakpoint at Xp11.2

CC within an ornithine-delta-aminotransferase OATL2 region, associated with
 CC the translocation t(X;18)(p11.2;q11.2) found in human synovial sarcomas,
 CC involving joining of the SYT gene (AAT11778) on chromosome-18 at 18q11.2
 CC to SSX2. The gene fusion is then transcribed to produce an SYT-SSX2
 CC (AAT11781) fusion transcript and translated into a fusion protein. SSX2
 CC may be distinguished from related sequence SSX1 (AAT11779) by digestion
 CC with SmaI and IspI. Primer AAT11784 has been used in polymerase chain
 CC reaction amplification of SSX2 sequences. Detection of the gene or its
 CC product in an abnormal location or as a fusion may be used in diagnosis
 CC of synovial sarcoma, using primers, probes, humanised antibodies,
 CC monoclonal antibodies, etc. Antisense oligonucleotides and antibodies may
 CC also be used therapeutically, and antitumour recombinant vaccines may be
 CC constructed e.g. in a vaccinia virus vector

SQ Sequence 766 BP; 229 A; 181 C; 200 G; 156 T; 0 U; 0 Other;

Query Match 86.2%; Score 496.6; DB 2; Length 766;
 Best Local Similarity 91.5%; Pred. No. 1.5e-146;
 Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 1 ATGAACGGAGACGACGCTTTGCAAGGAGACCCAGGGATGCTCAAAATATCAGAGAAG 60
 DB 92 ATGAACGGAGACGACGCTTTGCAAGGAGACCCAGGGATGCTCAAAATATCAGAGAAG 151
 QY 61 TTACGAAAGCCCTTCGATGATATGCAAAATCTCTTAAGAAAGAGTGGAAAGATG 120
 DB 152 ATCAAAAGCCCTTCGATGATATGCAAAATCTCTTAAGAAAGAGTGGAAAGATG 211
 QY 121 AAATCTCGGAGAAATCGTCTATGTATATGAAGCTAAATATGAGTCTATGACTAA 180
 DB 212 AAAGCCCTCGGAGAAATCTCTATGTATATGAAGAAAGATGATGAGGCTATGACTAA 271
 QY 181 CTAGGTTTCAAGTCCACCTCCACCTTCATCGCTAGTAAAGCGGCTCGAGCTCCAC 240
 DB 272 CTAGGTTTCAAGGCGCCCTCCACCTTCATGTGTATTAAGCGGCGGAAAGCTCCAG 331
 QY 241 GGAATGATTTTGGTAAACGATCAACACACAGGAATCAGGTTGAACGCTCTCAGATGCT 300
 DB 332 GGAATGATTTTGGTAAATGATGACCTAACCGTGGGAATCAGGTTGAACGCTCTCAGATGCT 391
 QY 301 TTGGGAGCTCCAGAGATCTTCCGAGATCATGCCNAGAGCCAGCAGAGAGAA 360
 DB 392 TTGGGAGCTCCAGAGATCTTCCGAGATCATGCCNAGAGCCAGCAGAGAGAA 451
 QY 361 AATGTTTGAAGGAAGTGCAGAGGATCTGGCCCAACAAATATGAGGAAACAGCTGTGC 420
 DB 452 AATGATTCGAGGAGTGCAGAGATCTGGCCCAACAAATATGAGGAAAGCTGTGC 511
 QY 421 CCCCAGGAAATCCAGTACCTTGGAGAGATTAACAAGACATCTGGACCCCAAGGGGG 480
 DB 512 CCCCAGGAAATCCAGTACCTTGGAGAGATTAACAAGACATCTGGACCCCAAGGGGG 571
 QY 481 AAACATGCTGGACCCAGACATCGGTGAGAGAAAGCAGCTGGTGTATGAAGATC 540
 DB 572 GACATGCTGGACCCAGACATCGGTGAGAGAAAGCAGCTGGTGTATGAAGATC 631
 QY 541 AGCGACCTCGAGAGATGACGAGTAACCTCCCTC 575
 DB 632 AGCGACCTCGAGAGATGACGAGTAACCTCCCTC 666

RESULT 6
 ABS73282
 ID ABS73282 standard; DNA; 766 BP.
 XX
 AC ABS73282;
 XX
 DT 04-DEC-2002 (first entry)
 XX
 DE DNA encoding human translocation (X; 18)(p11.2;q11.2) protein #2.
 XX
 KW Chromosome aberration; oncogenic fusion protein; cancer;
 KW proliferative disease; cellular protein isoform; heat shock protein 90;

KW HSP-90; rheumatoid arthritis; cancer; haematopoietic disorder;
 KW T cell lymphoma; B cell lymphoma; chronic myeloid leukaemia; CML;
 KW acute myeloid leukaemia; AML; chronic myelomonocytic leukaemia; CMML;
 KW acute lymphoblastic leukaemia; ALL; APL; NHL; solid tumour;
 KW papillary thyroid carcinoma; Ewing's sarcoma; melanoma; liposarcoma;
 KW rhabdomyosarcoma; synovial sarcoma; viral infection; gene; ds.
 XX Homo sapiens.
 OS WO200269900-A2.
 PN 12-SEP-2002.
 PD 01-MAR-2002; 2002WO-US0006518.
 XX 01-MAR-2001; 2001US-0272751P.
 XX (CONF-) CONFORMA THERAPEUTICS CORP.
 PA Fritz LC, Burrows PJ;
 PI WPI; 2002-698710/75.
 DR P-PSDB; ABG95081.
 XX
 PT Treating genetically-defined disease associated with chromosomal
 PT aberrations yielding oncogenic fusion proteins, e.g. cell proliferative
 PT diseases, involves administering an inhibitor of heat shock protein 90.
 XX Disclosure; Page 236; 389pp; English.

The invention describes a method of treating genetically-defined disease associated with chromosomal aberrations yielding oncogenic fusion proteins (I), treating cancerous cells containing (I) in a heterogeneous cell population, treating proliferative diseases associated with mutant protein or cellular protein isoforms (II) dependent on heat shock protein (HSP)-90, or selectively treating cells expressing (II) involving administering HSP90-inhibitor. The method is useful for treating genetically-defined disease with chromosomal aberration yielding oncogenic fusion protein, treating cancerous cells containing fusion protein in heterogeneous cell population, treating proliferative disease (e.g. rheumatoid arthritis or cancer) associated with mutant protein or cellular protein isoform dependent on heat shock protein (HSP)-90 (e.g. p53), or selectively treating cells expressing mutant protein or cellular protein isoform in a patient heterozygous for (II). The method is useful for treating a disease e.g. haematopoietic disorder such as T or B cell lymphoma, chronic myeloid leukaemia (CML), APL, ALL, NHL and CMML, or a disease characterised by a solid tumour such as papillary thyroid carcinoma, Ewing's sarcoma, melanoma, liposarcoma, rhabdomyosarcoma and synovial sarcoma. The method is also useful for treating viral infections. This represents the DNA sequence of a chromosome aberration

SQ Sequence 766 BP; 229 A; 181 C; 200 G; 156 T; 0 U; 0 Other;
 Query Match 86.2%; Score 496.6; DB 6; Length 766;
 Best Local Similarity 91.5%; Pred. No. 1.5e-146;
 Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 1 ATGAACGGAGACGACGCTTTGCAAGGAGACCCAGGGATGCTCAAAATATCAGAGAAG 60
 DB 92 ATGAACGGAGACGACGCTTTGCAAGGAGACCCAGGGATGCTCAAAATATCAGAGAAG 151
 QY 61 TTACGAAAGCCCTTCGATGATATGCAAAATCTCTTAAGAAAGAGTGGAAAGATG 120
 DB 152 ATCAAAAGCCCTTCGATGATATGCAAAATCTCTTAAGAAAGAGTGGAAAGATG 211
 QY 121 AAATCTCGGAGAAATCGTCTATGTATATGAAGCTAAATATGAGGCTATGACTAA 180
 DB 212 AAAGCCCTCGGAGAAATCTCTATGTATATGAAGAAAGATGATGAGGCTATGACTAA 271
 QY 181 CTAGGTTTCAAGGTCACCTCCACCTTCATGCTAGTAAAGCGGCTCGAGCTCCAC 240
 DB 272 CTAGGTTTCAAGGCGCCCTCCACCTTCATGTGTATTAAGCGGCGGAAAGCTCCAG 331

QY 241 GGGATGATTGGTAAACGATCAAAACACAGGAATCAGGTTGAACGTCCTCAGATGACT 300
 Db 332 GGGATGATTGGTAAACGATCAAAACACAGGAATCAGGTTGAACGTCCTCAGATGACT 391
 QY 301 TTCCGGCAGCTCCAGAGATCTTCCCGAGATCATGCCCAAGAGCCAGCAGAGAGAA 360
 Db 392 TTCCGGCAGCTCCAGAGATCTTCCCGAGATCATGCCCAAGAGCCAGCAGAGAGAA 451
 QY 361 AATGTTTGAAGGAAGTCCAGAGGATCTGGCCCAACAAATGATGGGAAAACAGTGTGC 420
 Db 452 AATGTTTGAAGGAAGTCCAGAGGATCTGGCCCAACAAATGATGGGAAAACAGTGTGC 511
 QY 421 CCCCCGGGAATCCAGTACCTTGGAGAGATTAACAGACATCTGGACCCAAAGGGGG 480
 Db 512 CCCCCGGGAATCCAGTACCTTGGAGAGATTAACAGACATCTGGACCCAAAGGGGG 571
 QY 481 AARATGCTCTGGACCCACAGACTGCGTGAGAGAAAGCAGTGGTGTATGAAGAGATC 540
 Db 572 GAACATGCTGGACCCACAGACTGCGTGAGAGAAAGCAGTGGTGTATGAAGAGATC 631
 QY 541 AGCGACCTGAGGAAGATGACGAGTAATCTCCCTC 575
 Db 632 AGCGACCTGAGGAAGATGACGAGTAATCTCCCTC 666

RESULT 7

AD54033
 ID AAD54033 standard; DNA; 766 BP.

XX AC AAD54033;
 XX

DT 17-JUN-2003 (first entry)
 XX

DE Human colon cancer-associated polypeptide gene, SSX-2.
 XX

KW Human; colon cancer-associated polypeptide; immune response; therapy;
 KW colon cancer; gene; ds.
 XX

OS Homo sapiens.
 XX

PN WO2002290986-A1.
 XX

PD 14-NOV-2002.
 XX

XX 02-MAY-2002; 2002WO-US013994.
 XX

XX 04-MAY-2001; 2001US-00849602.
 XX

XX (LUDW-) LUDWIG INST CANCER RES.
 XX (CORR) CORNELL RES FOUND INC.
 PA

XX Chen Y, Old LJ, Scanlan MJ, Stockert E;
 XX WPI; 2003-112003/10.
 XX

XX Diagnosing colon cancer in a subject comprises identifying colon cancer-
 PT associated polypeptides as antigens that elicit immune response in colon
 PT cancer.
 XX

XX Claim 1; Page 90-91; 122pp; English.
 XX

XX The invention relates to a method for diagnosing colon cancer in a
 CC subject which comprises identifying colon cancer-associated polypeptides
 CC as antigens that elicit immune response in colon cancer. The method is
 CC useful for diagnosing, determining onset, progression, or regression of
 CC colon cancer in a subject, or for selecting a course of treatment of a
 CC subject having or suspected of having colon cancer. The colon cancer-
 CC associated polypeptides are useful as markers for diagnosing colon
 CC cancer, and for following the course of treatment of colon cancer. The
 CC present sequence is human colon cancer-associated polypeptide gene
 XX

XX Sequence 766 BP; 229 A; 181 C; 200 G; 156 T; 0 U; 0 Other;
 SQ

Query Match 86.2%; Score 496.6; DB 7; Length 766;
 Best Local Similarity 91.5%; Pred. No. 1.5e-146;
 Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
 QY 1 ATGAACGGAGACGACGCTTTTGAAGGAGACCCAGAGATGATGCTCAAAATATCAGAGAAG 60
 Db 92 ATGAACGGAGACGACGCTTTTGAAGGAGACCCAGAGATGATGCTCAAAATATCAGAGAAG 151
 QY 61 TTACGAAGAGCCTTCGATGATATTGCCAAATATCTTCTTAAGAAAGAGTGGGAAAAGATG 120
 Db 152 ATCCAAAGAGCCTTCGATGATATTGCCAAATATCTTCTTAAGAAAGAGTGGGAAAAGATG 211
 QY 121 AATTCCTCGGAGAAAATCGTCTATGTATATCAAGCTAAACTATGAGGTCTAGCTATAA 180
 Db 212 AAGCCTCGGAGAAAATCTTCTATGTATATGAAGAAAAGTATGAGGTATGACTATAA 271
 QY 181 CTAGGTTTCAAGGTCACCTTCCACCTTTTCATCGGTAGTAAACGGGCTGCAGACTTCCAC 240
 Db 272 CTAGGTTTCAAGGTCACCTTCCACCTTTTCATCGGTAGTAAACGGGCTGCAGACTTCCAC 331
 QY 241 GGGATGATTGGTAAACGATCAAAACACAGGAATCAGGTTGAACGTCCTCAGATGACT 300
 Db 332 GGGATGATTGGTAAACGATCAAAACACAGGAATCAGGTTGAACGTCCTCAGATGACT 391
 QY 301 TTCCGGCAGCTCCAGAGATCTTCCCGAGATCATGCCCAAGAGCCAGCAGAGAGAA 360
 Db 392 TTCCGGCAGCTCCAGAGATCTTCCCGAGATCATGCCCAAGAGCCAGCAGAGAGAA 451
 QY 361 AATGTTTGAAGGAAGTCCAGAGGATCTGGCCCAACAAATGATGGGAAAACAGTGTGC 420
 Db 452 AATGTTTGAAGGAAGTCCAGAGGATCTGGCCCAACAAATGATGGGAAAACAGTGTGC 511
 QY 421 CCCCCGGGAATCCAGTACCTTGGAGAGATTAACAGACATCTGGACCCAAAGGGGG 480
 Db 512 CCCCCGGGAATCCAGTACCTTGGAGAGATTAACAGACATCTGGACCCAAAGGGGG 571
 QY 481 AARATGCTCTGGACCCACAGACTGCGTGAGAGAAAGCAGTGGTGTATGAAGAGATC 540
 Db 572 GAACATGCTGGACCCACAGACTGCGTGAGAGAAAGCAGTGGTGTATGAAGAGATC 631
 QY 541 AGCGACCTGAGGAAGATGACGAGTAATCTCCCTC 575
 Db 632 AGCGACCTGAGGAAGATGACGAGTAATCTCCCTC 666

RESULT 8

ABQ83844
 ID ABQ83844 standard; cDNA; 766 BP.

XX AC ABQ83844;
 XX

XX 03-FEB-2003 (first entry)
 XX

XX Human SSX-2 encoding cDNA SEQ ID NO:6.
 XX

XX Human; epitope; vaccine; immunotherapeutic; cytostatic; immunogenicity;
 KW T cell; gene; ss.
 XX

OS Homo sapiens.
 XX

XX WO200281646-A2.
 XX

XX 17-OCT-2002.
 XX

XX 04-APR-2002; 2002WO-US011101.
 XX

XX 06-APR-2001; 2001US-0282211P.
 PR 07-NOV-2001; 2001US-0337017P.
 PR 07-MAR-2002; 2002US-0363210P.
 XX

XX (CTL1-) CTL IMMUNOTHERAPIES CORP.
 PA
 XX Simard JLL, Diamond DC, Liu L, Xie Z;
 PI


```
XX DR WPI; 2003-067518/06.
XX P-PSDB; ABP741130.
XX
PT Novel epitopes useful as vaccines, comprises peptides or nucleic acid
PT encoding the peptides, that are useful epitopes of target-associated
PT antigens.
PS
PS Claim 1; Page 148; 352pp; English.
XX
CC The present invention describes an isolated epitope (I) and an epitope
CC cluster. Also described is a vaccine or immunotherapeutic composition
CC (VC) comprising (I). (I) has cytostatic activity. VC is useful for
CC treating an animal, by administering to an animal the vaccine or
CC immunotherapeutic composition. VC is also useful for evaluating
CC immunogenicity of a vaccine or immunotherapeutic composition, by
CC administering VC to an HLA-transgenic animal and evaluating
CC immunogenicity based on a characteristic of the animal, or by in vitro
CC primary stimulation of a T cell and evaluating immunogenicity. (I) is
CC useful for determining specific T cell frequency, by contacting T cells
CC with a MHC-peptide complex, and further comprises ELISPOT analysis,
CC limiting dilution analysis, flow cytometry, in situ hybridisation and/or
CC polymerase chain reaction (PCR). AB083843 to AB083858 and ABP74128 to
CC ABP741713 represent sequences used in the exemplification of the present
CC invention
XX
SQ Sequence 766 BP; 229 A; 181 C; 200 G; 156 T; 0 U; 0 Other;
Query Match 86.2%; Score 496.6; DB 7; Length 766;
Best Local Similarity 91.5%; Pred. No. 1.5e-146;
Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
QY 1 ATGAACGGAGACGACGCTTTGCAAGGAGACCCAGGATGATGCTCAATATCAGAGAAG 60
DB 92 ATGAACGGAGACGACGCTTTGCAAGGAGACCCAGGATGATGCTCAATATCAGAGAAG 151
QY 61 TTACGAAAGGCTTCGATGATATGCGCAATATCTCTAAGAAAGAGTGGGAAAGATG 120
DB 152 ATCCAAAGGCTTCGATGATATGCGCAATATCTCTAAGAAAGAGTGGGAAAGATG 211
QY 121 AAATCCTCGGAGAAATCGTCTATGTATATGAAGCTAACTATGAGGTGATGACTAAA 180
DB 212 AAAGCCTCGGAGAAATCTTCTATGTATATGAAGAAAGATGATGAGGTGATGACTAAA 271
QY 181 CTAGTTCGAGGTGACGCTCCACCTTTCATGTGTAGTAACCGGTGCGAGCTTCCAC 240
Query Match 86.2%; Score 496.6; DB 9; Length 766;
Best Local Similarity 91.5%; Pred. No. 1.5e-146;
Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
QY 1 ATGAACGGAGACGACGCTTTGCAAGGAGACCCAGGATGATGCTCAATATCAGAGAAG 60
DB 92 ATGAACGGAGACGACGCTTTGCAAGGAGACCCAGGATGATGCTCAATATCAGAGAAG 151
QY 61 TTACGAAAGGCTTCGATGATATGCGCAATATCTCTAAGAAAGAGTGGGAAAGATG 120
DB 152 ATCCAAAGGCTTCGATGATATGCGCAATATCTCTAAGAAAGAGTGGGAAAGATG 211
QY 121 AAATCCTCGGAGAAATCGTCTATGTATATGAAGCTAACTATGAGGTGATGACTAAA 180
DB 212 AAAGCCTCGGAGAAATCTTCTATGTATATGAAGAAAGATGATGAGGTGATGACTAAA 271
QY 181 CTAGTTCGAGGTGACGCTCCACCTTTCATGTGTAGTAACCGGTGCGAGCTTCCAC 240
Query Match 86.2%; Score 496.6; DB 9; Length 766;
Best Local Similarity 91.5%; Pred. No. 1.5e-146;
Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
QY 1 ATGAACGGAGACGACGCTTTGCAAGGAGACCCAGGATGATGCTCAATATCAGAGAAG 60
DB 92 ATGAACGGAGACGACGCTTTGCAAGGAGACCCAGGATGATGCTCAATATCAGAGAAG 151
QY 61 TTACGAAAGGCTTCGATGATATGCGCAATATCTCTAAGAAAGAGTGGGAAAGATG 120
DB 152 ATCCAAAGGCTTCGATGATATGCGCAATATCTCTAAGAAAGAGTGGGAAAGATG 211
QY 121 AAATCCTCGGAGAAATCGTCTATGTATATGAAGCTAACTATGAGGTGATGACTAAA 180
DB 212 AAAGCCTCGGAGAAATCTTCTATGTATATGAAGAAAGATGATGAGGTGATGACTAAA 271
QY 181 CTAGTTCGAGGTGACGCTCCACCTTTCATGTGTAGTAACCGGTGCGAGCTTCCAC 240
```

```
RESULT 9
ADC09570 ADC09570 standard; cDNA; 766 BP.
XX AC ADC09570;
XX XX 18-DEC-2003 (first entry)
XX DE SSX-2 cDNA #SEQ ID 6.
XX KW Epitope; immunological; vaccine;
XX KW major histocompatibility complex class I; MHC class I; cancer;
XX KW immunisation; ss.
XX OS Unidentified.
XX PN WO2003008537-A2.
XX PD 30-JAN-2003.
XX PF 29-MAR-2002; 2002WO-US010189.
XX PR 06-APR-2001; 2001US-0282211P.
XX PR 07-NOV-2001; 2001US-0337017P.
XX PR 07-MAR-2002; 2002US-0383210P.
XX PA (CTLI-) CTL IMMUNOTHERAPIES CORP.
XX PI Simard JUL, Diamond DC, Liu L, Xie Z;
XX DR WPI; 2003-248010/24.
XX PT Epitope having high affinity for major histocompatibility complex class I
XX PT useful for treating an animal, evaluating immunogenicity of a vaccine or
XX PT therapeutic composition and for diagnosing a disease.
XX PS Claim 1; SEQ ID NO 6; 239pp; English.
XX CC The invention relates to an isolated epitope polypeptide that has high
XX CC affinity for major histocompatibility complex (MHC) class I, and an
XX CC epitope cluster comprising the polypeptide. Also disclosed is a vaccine
XX CC or immunotherapeutic composition containing an epitope of the invention.
XX CC Compositions of the invention may be used in the treatment of cancer. The
XX CC method can be combined with a radiation therapy, chemotherapy,
XX CC biochemotherapy or surgery. The composition is also useful for evaluating
XX CC immunogenicity of a vaccine or immunotherapeutic compound. Multimeric MHC
XX CC peptide complexes of the invention are useful for determining specific T
XX CC cell frequency. This method is useful for evaluating immunological
XX CC response, by performing the method prior to and subsequent to an
XX CC immunisation step. Compositions of the invention are useful for
XX CC diagnosing a disease. The current sequence represents an epitope of the
XX CC invention with high affinity for MHC class I.
XX SQ Sequence 766 BP; 229 A; 181 C; 200 G; 156 T; 0 U; 0 Other;
Query Match 86.2%; Score 496.6; DB 9; Length 766;
Best Local Similarity 91.5%; Pred. No. 1.5e-146;
Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
QY 1 ATGAACGGAGACGACGCTTTGCAAGGAGACCCAGGATGATGCTCAATATCAGAGAAG 60
DB 92 ATGAACGGAGACGACGCTTTGCAAGGAGACCCAGGATGATGCTCAATATCAGAGAAG 151
QY 61 TTACGAAAGGCTTCGATGATATGCGCAATATCTCTAAGAAAGAGTGGGAAAGATG 120
DB 152 ATCCAAAGGCTTCGATGATATGCGCAATATCTCTAAGAAAGAGTGGGAAAGATG 211
QY 121 AAATCCTCGGAGAAATCGTCTATGTATATGAAGCTAACTATGAGGTGATGACTAAA 180
DB 212 AAAGCCTCGGAGAAATCTTCTATGTATATGAAGAAAGATGATGAGGTGATGACTAAA 271
QY 181 CTAGTTCGAGGTGACGCTCCACCTTTCATGTGTAGTAACCGGTGCGAGCTTCCAC 240
```

Db 272 CTAGGTTTCAAGGCCACCTCCACCTTTTCATGTGTAATAAAGCGGCCGAAAGACTTCCAG 331
 Qy 241 GGAATGATTTTGTAAACGATCGAAACCCACAGGAATCAGGTTGAACGTCTCTCAGATGACT 300
 Db 332 GGAATGATTTTGTAAATGATACCTTAACCGTGGGAATCAGGTTGAACGTCTCTCAGATGACT 391
 Qy 301 TTCCGGACGCTCCAGAGAAATCTTCCCGAAGATCATGCCAAGAGCCACAGAGAGAGAA 360
 Db 392 TTCCGGACGCTCCAGAGAAATCTTCCCGAAGATCATGCCAAGAGCCACAGAGAGAGAA 451
 Qy 361 AATGTTTGAAGAAAGTCCAGAGGATCTTGGCCCAAAATATGATGGGAAAACAGCTGTGC 420
 Db 452 AATGTTTGAAGAAAGTCCAGAGGATCTTGGCCCAAAATATGATGGGAAAACAGCTGTGC 511
 Qy 421 CCCCCGGGAATCAAGTACCTTGGAGAGATTTAACAAGACATCTCGACCCAAAAGGGGG 480
 Db 512 CCCCCGGGAATCAAGTACCTTGGAGAGATTTAACAAGACATCTCGACCCAAAAGGGGG 571
 Qy 481 AAACATGCTCGGACCCACAGACTGCGTGAGAGAAAGCAGCTGGTGGTTTATGAAGAGATC 540
 Db 572 GAACATGCTCGGACCCACAGACTGCGTGAGAGAAAGCAGCTGGTGGTTTATGAAGAGATC 631
 Qy 541 AGGACCTCGAGGAAGATGACGAGTAAGTAACTCCCTC 575
 Db 632 AGGACCTCGAGGAAGATGACGAGTAAGTAACTCCCTC 666

RESULT 10

ADD25523

ID ADD25523 standard; DNA; 766 BP.

XX AC

XX ADD25523;

XX XX

XX 15-JAN-2004 (first entry)

XX Binding domain-immunoglobulin fusion protein-associated DNA #46.

XX ds;

KW ds; Binding domain; immunoglobulin; fusion protein; cytostatic;
 KW antiarthritic; immunosuppressive; antidiabetic; antithyroid;
 KW neuroprotective; hinge region; immunoglobulin heavy chain;
 KW CH2 constant region; CH3 constant region; IgG1;
 KW antibody dependent cell-mediated cytotoxicity; ADCC; complement fixation;
 KW malignant condition; B-cell disorder; melanoma; carcinoma; sarcoma;
 KW rheumatoid arthritis; myasthenia gravis; Grave's disease;
 KW type I diabetes mellitus; multiple sclerosis; autoimmune disease.

XX Unidentified.

XX OS

XX US2003118592-A1.

XX PN

XX 26-JUN-2003.

XX PD

XX 25-JUL-2002; 2002US-00207655.

XX PF

XX 17-JAN-2001; 2001US-0367358P.

XX PR

XX 17-JAN-2002; 2002US-00053530.

XX PR

XX 03-JUN-2002; 2002US-0385691P.

XX PR

XX (GENE-) GENE-CRAFT INC.

XX PA

XX Ledbetter JA, Hayden-Ledbetter MS, Thompson PA;

XX PI

XX WPI; 2003-801317/75.

XX DR

XX New binding domain-immunoglobulin fusion protein, useful for treating a
 PT subject having or suspected of having a malignant condition or a B-cell
 PT disorder, e.g. melanoma, Grave's disease or autoimmune disease.

XX XX

XX PS

XX Disclosure; SEQ ID NO 84; 157pp; English.

XX CC

XX Unidentified

XX XX

XX Sequence 766 BP; 229 A; 181 C; 200 G; 156 T; 0 U; 0 Other;

XX SQ

Query Match 86.2%; Score 496.6; DB 9; Length 766;
 Best Local Similarity 91.5%; Pred. No. 1.5e-146;
 Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

Qy 1 ATGAACGGAGACGACGCTTTGCAAGAGACCCAGGATGATCTCAAAATATCAGAGAAG 60
 Db 92 ATGAACGGAGACGACGCTTTGCAAGAGACCCAGGATGATCTCAAAATATCAGAGAAG 151
 Qy 61 TTACGAAGGCTTTGATGATATTGCGCAATACTTCTTAAGAAAGAGTGGGAAAAGATG 120
 Db 152 ATCCAAAGGCTTTGATGATATTGCGCAATACTTCTTAAGAAAGAGTGGGAAAAGATG 211
 Qy 121 AAATCTCTGGAGAAATCGTCTATGTATATGAAGCTAACTATGAGTCTAGCTATAA 180
 Db 212 AAAGCTCTGGAGAAATCGTCTATGTATATGAAGAAAGTATGAGGCTATGACTAAA 271
 Qy 181 CTAGGTTTCAAGGTCACCTCCACCTTTTCATGCTAGTAAACGGGCTGCAGACTTCCAC 240
 Db 272 CTAGGTTTCAAGGTCACCTCCACCTTTTCATGCTAGTAAACGGGCTGCAGACTTCCAG 331
 Qy 241 GGGAAATGATTTGGTAACGATCGAAACCCACAGGAATCAGGTTGAACGTCTCAGATGACT 300
 Db 332 GGGAAATGATTTGGTAACGATCGAAACCCACAGGAATCAGGTTGAACGTCTCAGATGACT 391
 Qy 301 TTCCGACGCTCCAGAGAAATCTTCCCGAAGATCATGCCAAGAGCCACAGAGGAAGAA 360
 Db 392 TTCCGACGCTCCAGAGAAATCTTCCCGAAGATCATGCCAAGAGCCACAGAGGAAGGA 451
 Qy 361 AATGTTTGAAGAAAGTCCAGAGGATCTTGGCCCAAAATATGATGGGAAAACAGCTGTGC 420
 Db 452 AATGTTTGAAGAAAGTCCAGAGGATCTTGGCCCAAAATATGATGGGAAAACAGCTGTGC 511
 Qy 421 CCCCCGGGAATCCAAGTACCTTGGAGAGATTAACAAGACATCTGGACCCAAAAGGGGG 480
 Db 512 CCCCCGGGAATCCAAGTACCTTGGAGAGATTAACAAGACATCTGGACCCAAAAGGGGG 571
 Qy 481 AAACATGCTCGACCCACAGACTGCGTGAGAGAAAGCAGCTGGTGGTTTATGAAGAGATC 540
 Db 572 GAACATGCTCGACCCACAGACTGCGTGAGAGAAAGCAGCTGGTGGTTTATGAAGAGATC 631
 Qy 541 AGGACCTCGAGGAAGATGACGAGTAAGTAACTCCCTC 575
 Db 632 AGGACCTCGAGGAAGATGACGAGTAAGTAACTCCCTC 666

RESULT 11

AAT47748

ID AAT47748 standard; cDNA; 931 BP.

XX AC

XX AAT47748;

XX DT

XX 05-JUN-1997 (first entry)

XX DE

XX Melanoma cell antigen HOM-MEL 40 cDNA clone.

XX XX

XX Melanoma; tumour antigen; HOM-MEL 40; serological fishing; vaccine; ds.

XX KW

XX Homo sapiens.

XX OS

XX WO9640209-A1.

XX PN

XX 19-DEC-1996.

XX PD

XX 07-JUN-1996; 96WO-US009726.

XX PF

XX 07-JUN-1995; 95US-00479328.

XX PR

XX 03-JAN-1996; 96US-00580980.

XX PR

XX 10-MAY-1996; 96US-00644116.

XX XX

XX (LUDW-) LUDWIG INST CANCER RES.

XX PA

XX Pfreundschuh M, Rammensee H;

XX PI

XX WPI; 1997-051896/05.
 XX Novel method of serological fishing - allows isolation of molecules (esp.
 PT antigens) associated with pathological conditions, e.g. Hodgkin's
 PT disease, cancer or auto-immune disease.
 XX
 XX Claim 31; Page 29; 44pp; English.
 XX
 CC A cDNA clone (AA747748) codes for a novel melanoma cell antigen
 CC designated HOM-MEL 40. It was isolated using a novel method in which a
 CC cDNA library prep'd. from malignant melanoma and expressed in E. coli
 CC cells. Lysates of the host cells were screened with sera that had been
 CC treated to remove interfering binding partners. This involved contacting
 CC the sample with lysates of untransfected host cells and with host cells
 CC transformed with the same vector (phage lambda) used to make the cDNA
 CC library. The method, termed serological fishing, can be used to detect
 CC antigens in human tissues, esp. tumour cells, which are useful in the
 CC molecular diagnosis of diseases and/or for immunotherapy and gene therapy
 CC of infectious, autoimmune and malignant diseases (see also AA747747-49).
 CC HOM-MEL 40 is expressed in 50% of melanomas, 20% of prostate cancers, 20%
 CC of gastric cancers, 26% of colorectal cancers, 12% of lung cancers and
 CC 20% of breast cancers, but not in normal tissue. HLA-A2 positive tumour
 CC cells present a nonamer (see also AA09449-52) derived from HOM-MEL,
 CC suggesting that HOM-MEL 40-specific vaccines, useful in inducing
 CC cytotoxic T lymphocytes, are possible
 XX
 SQ Sequence 931 BP; 274 A; 207 C; 231 G; 219 T; 0 U; 0 Other;

Query Match 86.2%; Score 496.6; DB 2; Length 931;
 Best Local Similarity 91.5%; Pred. No. 1.6e-146;
 Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
 1 ATGAACGAGACGACGCTTTGCAAGGAGACCCAGGAGTATGCTCAAAATATCAGAGAAG 60
 99 ATGAACGAGACGACGCTTTGCAAGGAGACCCAGGAGTATGCTCAAAATATCAGAGAAG 158
 61 TTACGAAAGGCTTCGATGATATTCGCAAAATATCTTCTAAGAAAGAGTGGGAAAGATG 120
 159 ATCCAAAGGCTTCGATGATATTCGCAAAATATCTTCTAAGAAAGAGTGGGAAAGATG 218
 121 AAATCTCGGAGAAATTCGTCTATGTATATGAAGCTAACTATAGGTGATGACTTAA 180
 219 AAAGCTCGGAGAAATTCGTCTATGTATATGAAGAAAGTATGAGGTGATGACTTAA 278
 181 CTAGGTTTCAAGGTCACCTCCACCTTTCATGCGTAGTAAACGGGTGACAGCTTCCAC 240
 Query Match 86.2%; Score 496.6; DB 2; Length 931;
 Best Local Similarity 91.5%; Pred. No. 1.6e-146;
 Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
 1 ATGAACGAGACGACGCTTTGCAAGGAGACCCAGGAGTATGCTCAAAATATCAGAGAAG 60
 99 ATGAACGAGACGACGCTTTGCAAGGAGACCCAGGAGTATGCTCAAAATATCAGAGAAG 158
 61 TTACGAAAGGCTTCGATGATATTCGCAAAATATCTTCTAAGAAAGAGTGGGAAAGATG 120
 159 ATCCAAAGGCTTCGATGATATTCGCAAAATATCTTCTAAGAAAGAGTGGGAAAGATG 218
 121 AAATCTCGGAGAAATTCGTCTATGTATATGAAGCTAACTATAGGTGATGACTTAA 180
 219 AAAGCTCGGAGAAATTCGTCTATGTATATGAAGAAAGTATGAGGTGATGACTTAA 278
 181 CTAGGTTTCAAGGTCACCTCCACCTTTCATGCGTAGTAAACGGGTGACAGCTTCCAC 240

RESULT 12
 AAV04267
 ID AAV04267 standard; cDNA; 931 BP.
 AC AAV04267;
 XX
 XX 22-JUN-1998 (first entry)
 XX
 XX Melanoma antigen clone HOM-MEL-40.
 XX
 XX Melanoma; marker; antigen; diagnosis; serological fishing; human;
 KW HOM-MEL-40; ss.
 XX
 XX Homo sapiens.
 XX
 XX WO9748721-A1.
 XX
 XX 24-DEC-1997.
 XX
 XX 23-JUN-1997; 97WO-US010926.
 XX
 XX 21-JUN-1996; 96US-00668128.
 XX
 XX (LUDW-) LUDWIG INST CANCER RES.
 PA
 XX Pfreundschuh M;
 XX
 XX WPI; 1998-063074/06.
 XX
 XX Nucleic acid and derived protein are markers for Hodgkin's disease - used
 PT in identifying immuno:reactive markers of disease.
 XX
 XX Example 7; Page 30-31; 47pp; English.

PS
 CC Clone HOM-MEL-40 was isolated from a human malignant melanoma cDNA
 CC library using a method, designated serological fishing, designed to
 CC identify immunoreactive markers of disease. In this method, cells
 CC characteristic of a disease are used to prepare a cDNA library for
 CC transformation of eukaryotic or prokaryotic cells, and the cells grown to
 CC express proteins. Patient serum is incubated with the cells used to
 CC prepare the library, but not transfected, to remove any components
 CC reactive with these cells, then the stripped sample is similarly treated
 CC with cells carrying the empty vector. The twice-stripped sample is
 CC incubated with lysate of the library cells, so that specific components
 CC in the sample may bind to the expression protein. Proteins that form
 CC immune complexes are identified as disease markers. In the case of HOM-
 CC MEL-40, the new melanoma associated antigen is strongly expressed in
 CC melanoma, but not healthy tissues. The deduced amino acid sequence
 CC includes 3 tumour-associated peptides (see AA41587-89) that bind to HLA-
 CC A2.1. The serological fishing method was also used to identify a claimed
 CC marker (see AAV04262) of Hodgkin's disease
 XX

SQ Sequence 931 BP; 274 A; 207 C; 231 G; 219 T; 0 U; 0 Other;
 Query Match 86.2%; Score 496.6; DB 2; Length 931;
 Best Local Similarity 91.5%; Pred. No. 1.6e-146;
 Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
 1 ATGAACGAGACGACGCTTTGCAAGGAGACCCAGGAGTATGCTCAAAATATCAGAGAAG 60
 99 ATGAACGAGACGACGCTTTGCAAGGAGACCCAGGAGTATGCTCAAAATATCAGAGAAG 158
 61 TTACGAAAGGCTTCGATGATATTCGCAAAATATCTTCTAAGAAAGAGTGGGAAAGATG 120
 159 ATCCAAAGGCTTCGATGATATTCGCAAAATATCTTCTAAGAAAGAGTGGGAAAGATG 218
 121 AAATCTCGGAGAAATTCGTCTATGTATATGAAGCTAACTATAGGTGATGACTTAA 180
 219 AAAGCTCGGAGAAATTCGTCTATGTATATGAAGAAAGTATGAGGTGATGACTTAA 278
 181 CTAGGTTTCAAGGTCACCTCCACCTTTCATGCGTAGTAAACGGGTGACAGCTTCCAC 240

Db 279 CTAGTTTCAAGCCACCTCCACCTTTTCATGTGTAATAAACGGGCCGAGACTTCAG 338
 Qy 241 GGAATGATTTGGTAACGATCGAAACACAGGAATCAGGTTGAACGTCCTCAGATGACT 300
 Db 339 GGAATGATTTGGTAATGATACCTTAACCGTGGGAATCAGGTTGAACGTCCTCAGATGACT 398
 Qy 301 TTGGGAGCTCCAGAGATCTTCCCGAGATCATGCCCAAGAGCCAGAGGAGGAA 360
 Db 399 TTGGGAGCTCCAGGATCTCCCGAAGATCATGCCCAAGAGCCAGAGGAGGAA 458
 Qy 361 AATGTTTGAAGGAATGCGAGGATCTGCGCCCAAAATGATGGGAAACAGCTGTGC 420
 Db 459 AATGTTTGGAGGATGCGAGGATCTGCGCCCAAAATGATGGGAAACAGCTGTGC 518
 Qy 421 CCCCCGGGAATCCAGTACCTTGGAGAGATTAACAAGACATCTGAGCCCAAGGGG 480
 Db 519 CCCCCGGGAATCCAGTACCTTGGAGAGATTAACAAGACATCTGAGCCCAAGGGG 578
 Qy 481 AAACATGCTGAGCCACACAGATGCGTGAGAGAAAGCAGCTGCTGTTTATGAAGATC 540
 Db 579 GAACATGCTGAGCCACACAGATGCGTGAGAGAAAGCAGCTGCTGTTTATGAAGATC 638
 Qy 541 AGCGACCTGAGGAGATGAGTAACTCCCTC 575
 Db 639 AGCGACCTGAGGAGATGAGTAACTCCCTC 673

RESULT 13

ABL64122
 ID ABL64122 standard; DNA; 1309 BP.

XX ABL64122;

XX 15-MAY-2002 (first entry)

XX Breast cancer related gene sequence SEQ ID NO:2459.

XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
 XX stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
 XX cytosolic; gene therapy; antineoplastic; Wilms' tumour; adenocarcinoma;
 XX gene; ds.

XX Homo sapiens.

XX WO200194629-A2.

XX 13-DEC-2001.

XX 30-MAY-2001; 2001WO-US010838.

XX 05-JUN-2000; 2000US-0209473P.

XX 05-JUN-2000; 2000US-0209531P.

XX 18-SEP-2000; 2000US-0233133P.

XX 18-SEP-2000; 2000US-0233617P.

XX 20-SEP-2000; 2000US-0234009P.

XX 20-SEP-2000; 2000US-0234034P.

XX 22-SEP-2000; 2000US-0234052P.

XX 22-SEP-2000; 2000US-0234509P.

XX 22-SEP-2000; 2000US-0234567P.

XX 25-SEP-2000; 2000US-0234923P.

XX 25-SEP-2000; 2000US-0234924P.

XX 25-SEP-2000; 2000US-0235077P.

XX 25-SEP-2000; 2000US-0235082P.

XX 25-SEP-2000; 2000US-0235134P.

XX 25-SEP-2000; 2000US-0235280P.

XX 26-SEP-2000; 2000US-0235637P.

XX 27-SEP-2000; 2000US-0235638P.

XX 27-SEP-2000; 2000US-0235711P.

XX 27-SEP-2000; 2000US-0235720P.

XX 27-SEP-2000; 2000US-0235840P.

XX 27-SEP-2000; 2000US-0235863P.

XX 28-SEP-2000; 2000US-0236028P.

XX 28-SEP-2000; 2000US-0236032P.

PR 28-SEP-2000; 2000US-0236033P.
 PR 28-SEP-2000; 2000US-0236034P.
 PR 28-SEP-2000; 2000US-0236109P.
 PR 28-SEP-2000; 2000US-0236111P.
 PR 29-SEP-2000; 2000US-0236842P.
 PR 29-SEP-2000; 2000US-0236891P.
 PR 02-OCT-2000; 2000US-0237172P.
 PR 02-OCT-2000; 2000US-0237173P.
 PR 02-OCT-2000; 2000US-0237278P.
 PR 02-OCT-2000; 2000US-0237294P.
 PR 02-OCT-2000; 2000US-0237316P.
 PR 03-OCT-2000; 2000US-0237425P.
 PR 03-OCT-2000; 2000US-0237598P.
 PR 03-OCT-2000; 2000US-0237604P.
 PR 03-OCT-2000; 2000US-0237606P.
 PR 03-OCT-2000; 2000US-0237608P.
 PR 01-NOV-2000; 2000US-0244867P.
 PR 01-NOV-2000; 2000US-0245084P.
 PA (AVAL-) AVALON PHARM.
 XX
 XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
 PI Soppet DR, Weaver Z;
 XX
 DR WPI; 2002-188264/24.
 XX
 PT Screening for anti-neoplastic agent involves exposing cells to a chemical
 PT agent to be tested for anti-neoplastic activity, and determining a change
 PT in expression of a gene of a signature gene set.
 XX
 PS Claim 1; SEQ ID NO 2459; 44pp; English.
 XX
 CC The present invention describes a method (M1) for screening for an anti-
 CC neoplastic agent. The method involves exposing cells to a chemical agent
 CC to be tested for anti-neoplastic activity, determining a change in
 CC expression of at least one gene (I) of a signature gene set, where (I)
 CC comprises a sequence (S) selected from 847 sequences (given in ABL61664
 CC to ABL70110), or is at least 95% identical to (S), where a change in
 CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
 CC activity and can be used in gene therapy. M1 can be used for screening an
 CC anti-neoplastic agent, and can be used for producing a product which is
 CC the data collected with respect to the anti-neoplastic agent as a result
 CC of M1, and the data is sufficient to convey the chemical structure and/or
 CC properties of the agent. M1 can be used in the treatment of cancer such
 CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
 CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
 CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
 CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilms'
 CC tumour
 XX
 SQ Sequence 1309 BP; 385 A; 329 C; 281 G; 309 T; 0 U; 5 Other;
 Query Match 86.2%; Score 496.6; DB 6; Length 1309;
 Best Local Similarity 91.5%; Pred. No. 1.9e-146;
 Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
 Qy 1 ATGACGAGACGACGCGCTTGCAGGAGACCCAGGATGATGCTCAATATCAGAGAAG 60
 Db 99 ATGACGAGACGACGCGCTTGCAGGAGACCCAGGATGATGCTCAATATCAGAGAAG 158
 Qy 61 TTACGAAAGGCGCTTCGATGATATTGCCAAATACCTCTCTAAGAAAGAGTGGGAAAGATG 120
 Db 159 ATCCAAAGGCGCTTCGATGATATTGCCAAATACCTCTCTAAGAAAGAGTGGGAAAGATG 218
 Qy 121 AAATCCTCGGAGAAATCGTCTATGTATATGAGCTAACTATGAGGTCATGACTACATAA 180
 Db 219 AAAGCCTCGGAGAAATCTTCTATGTATATGAGGAGAAAGTATGAGGTCATGACTACATAA 278
 Qy 181 CTAGGTTTCAAGGTCACCTCCACCTTTCATGCTAGTAGTAAACCGGCTGCAGACTTCCAC 240
 Db 279 CTAGGTTTCAAGGTCACCTCCACCTTTCATGCTAGTAGTAAACCGGCTGCAGACTTCCAG 338

QY 241 GGGATGATTTGGTAAACGATCGAATCCAAACACAGGATCAGGTTGACGTCCTCAGATGACT 300
Db |||||
QY 339 GGGATGATTTGGTAAACGATCGAATCCAAACACAGGATCAGGTTGACGTCCTCAGATGACT 398
Db |||||
QY 301 TTGCGCAGCTCCAGAGATCTTCCGGAAGATCATGCCCCAAGAACCCAGAGAGAGAA 360
Db |||||
QY 399 TTGCGCAGCTCCAGAGATCTTCCGGAAGATCATGCCCCAAGAACCCAGAGAGAGAA 458
QY 361 AATGTTTGAAGGAATGCGAGAGCATCTGCCCCAAGAACCCAGAGAGAGAA 420
Db |||||
QY 459 AATGATTGGAAGGAATGCGAGAGCATCTGCCCCAAGAACCCAGAGAGAGAA 518
Db |||||
QY 421 CCCCCGGGAAATCCAAATGACCTTGAGAGAGATTAAAGACATCTGGACCCCAAGAGGGG 480
Db |||||
QY 519 CCCCCGGGAAATCCAAATGACCTTGAGAGAGATTAAAGACATCTGGACCCCAAGAGGGG 578
QY 481 AAACATGCTGACCCAGAGATGCTGAGAGAGAGAGAGAGAGAGAGAGAGATC 540
Db |||||
QY 579 GAACATGCTGACCCAGAGATGCTGAGAGAGAGAGAGAGAGAGAGAGAGATC 638
QY 541 AGCGACCTGAGGAAGATGACGAGTAACCTCCCTC 575
Db |||||
QY 639 AGCGACCTGAGGAAGATGACGAGTAACCTCCCTC 673

RESULT 14
ID ABL63730
XX ABL63730 standard; DNA; 1309 BP.
AC ABL63730;
XX
DT 15-MAY-2002 (first entry)
XX
DE Breast cancer related gene sequence SEQ ID NO:2067.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US010838.
XX
PR 05-JUN-2000; 2000US-0209473P.
PR 05-JUN-2000; 2000US-0209531P.
PR 18-SEP-2000; 2000US-0233133P.
PR 18-SEP-2000; 2000US-0233617P.
PR 20-SEP-2000; 2000US-0234009P.
PR 20-SEP-2000; 2000US-0234034P.
PR 22-SEP-2000; 2000US-0234052P.
PR 22-SEP-2000; 2000US-0234059P.
PR 22-SEP-2000; 2000US-0234567P.
PR 25-SEP-2000; 2000US-0234923P.
PR 25-SEP-2000; 2000US-0234924P.
PR 25-SEP-2000; 2000US-0235077P.
PR 25-SEP-2000; 2000US-0235082P.
PR 25-SEP-2000; 2000US-0235134P.
PR 25-SEP-2000; 2000US-0235280P.
PR 26-SEP-2000; 2000US-0235637P.
PR 26-SEP-2000; 2000US-0235638P.
PR 27-SEP-2000; 2000US-0235711P.
PR 27-SEP-2000; 2000US-0235720P.
PR 27-SEP-2000; 2000US-0235840P.
PR 28-SEP-2000; 2000US-0235863P.
PR 28-SEP-2000; 2000US-0236028P.
PR 28-SEP-2000; 2000US-0236032P.
PR 28-SEP-2000; 2000US-0236033P.
PR 28-SEP-2000; 2000US-0236034P.

PR 28-SEP-2000; 2000US-0236109P.
PR 28-SEP-2000; 2000US-0236111P.
PR 29-SEP-2000; 2000US-0236842P.
PR 29-SEP-2000; 2000US-0236891P.
PR 02-OCT-2000; 2000US-0237172P.
PR 02-OCT-2000; 2000US-0237173P.
PR 02-OCT-2000; 2000US-0237278P.
PR 02-OCT-2000; 2000US-0237294P.
PR 02-OCT-2000; 2000US-0237295P.
PR 02-OCT-2000; 2000US-0237316P.
PR 03-OCT-2000; 2000US-0237425P.
PR 03-OCT-2000; 2000US-0237598P.
PR 03-OCT-2000; 2000US-0237604P.
PR 03-OCT-2000; 2000US-0237606P.
PR 03-OCT-2000; 2000US-0237608P.
PR 01-NOV-2000; 2000US-0244867P.
PR 01-NOV-2000; 2000US-0245084P.
XX
PA (AVAL-) AVALON PHARM.
XX
PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX
XX WPI; 2002-188264/24.
XX
PT Screening for anti-neoplastic agent involves exposing cells to a chemical
PT agent to be tested for anti-neoplastic activity, and determining a change
PT in expression of a gene of a signature gene set.
XX
PS Claim 1; SEQ ID NO 2067; 44pp; English.
XX
CC The present invention describes a method (M1) for screening for an anti-
CC neoplastic agent. The method involves exposing cells to a chemical agent
CC to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences given in ABL61664
CC to ABL70110, or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening an
CC anti-neoplastic agent, and can be used for producing a product which is
CC the data collected with respect to the anti-neoplastic agent as a result
CC of M1, and the data is sufficient to convey the chemical structure and/or
CC properties of the agent. M1 can be used in the treatment of cancer such
CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's
CC tumour
XX
SQ Sequence 1309 BP; 385 A; 329 C; 281 G; 309 T; 0 U; 5 Other;
Query Match 86.2%; Score 496.6; DB 6; Length 1309;
Best Local Similarity 91.5%; Pred. NO. 1.9e-146;
Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
QY 1 ATGAACGGAGACAGACGCTTTGCAAGGAGACACCGAGATGCTCAATAATCAGAGAAG 60
Db 99 ATGAACGGAGACAGACGCTTTGCAAGGAGACACCGAGATGCTCAATAATCAGAGAAG 158
QY 61 TTACGAAAGGCTTCGATGATTTGCCAAATACTTCTTAAGAAAGAGTGGGAAAAGATG 120
Db 159 ATCCAAAAGGCTTCGATGATTTGCCAAATACTTCTTAAGAAAGAGTGGGAAAAGATG 218
QY 121 AAATCTCCGGAGAAATCGTCTATGTATATGAGCTAAACTATGAGTCTGACTAA 180
Db 219 AAAGCTTCGGAGAAATCTTCTATGTATATGAGAGAAAGATGATGAGGTATGACTAA 278
QY 181 CTAGGTTTCAAGGTCACCTCCACCTTCATGCGTAGTAAACGGGCTCGAGACTTCCAC 240
Db 279 CTAGGTTTCAAGGCCACCTCCACCTTCATGCTGTATATAAACGGGCCGAGACTTCCAG 338
QY 241 GGGAAATGATTTGGTAAACGATCGAATCCAAACACAGGATCAGGTTGACGTCCTCAGATGACT 300

Db 339 GGGAAATGTTGGATATGATACCTAACCGTGGGAATCAGGTTGAACGTCCTCAGATGACT 398
 Qy 301 TTCGGCAGCTCCAGAGAAATCTTCCGGAAGATCATGCCCAAGAACGACGAGAGAGAA 360
 Db 399 TTCGGCAGCTCCAGAGAAATCTTCCGGAAGATCATGCCCAAGAACGACGAGAGAGAA 458
 Qy 361 AATGGTTTGAAGAGTCCAGAGGATCTGCGCCCAAAATGATGGGAAACAGCTGTGC 420
 Db 459 AATGATTCGGAGGAAGTCCAGAGGATCTGCGCCCAAAATGATGGGAAACAGCTGTGC 518
 Qy 421 CCCCCGGGAATCCAAAGTACCTTGGAGAGATTAACAAGACATCTGACCCCAAAAGGGGG 480
 Db 519 CCCCCGGGAATCCAAAGTACCTTGGAGAGATTAACAAGACATCTGACCCCAAAAGGGGG 578
 Qy 481 AATAGCTCCGAGCCACAGACTGCGTGAGAGAAACAGCTGTGTGTTATGAAGAGATC 540
 Db 579 GAACATGCTCGGACCCACAGACTGCGTGAGAGAAACAGCTGTGTGTTATGAAGAGATC 638
 Qy 541 AGCGACCTGAGGAAGATGACGAGTAACCTCCCTC 575
 Db 639 AGCGACCTGAGGAAGATGACGAGTAACCTCCCTC 673

RESULT 15

AAS80402
 ID AAS80402 standard; cdna; 921 BP.

AC AAS80402;

XX 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #16206.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US008631.

XX 31-MAR-2000; 2000US-00540217.

XX 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX P-PSDB; ABG16215.

XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.

XX Claim 1; SEQ ID NO 16206; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The

CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
 CC coding sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX

Qy Sequence 921 BP; 267 A; 205 C; 233 G; 216 T; 0 U; 0 Other;

Query Match 84.68; Score 487.2; DB 5; Length 921;

Best Local Similarity 91.3%; Pred. No. 1.6e-143;

Matches 528; Conservative 0; Mismatches 48; Indels 2; Gaps 1;

Qy 1 ATGAACGGAGACGACGCTTTTGAAGGAGACCCAGGAGATGATCTCAAAATATCAGAGAAG 60

Db 54 ATGAACGGAGATGACACCTTTTGAAGGAGACCCAGGCTGGTCTCAATACACAGAGAAG 113

Qy 61 TTACGAAGGCTTCGATGATATTCCTCAATATCTTCTTAAGAAAGAGTGGGAAGATG 120

Db 114 ATCAAAAGGCTTCGATGATATTCCTCAATATCTTCTTAAGAAAGAGTGGGAAGATG 173

Qy 121 AAATCCTCGGAGAAATCGTCTATGTATGAAGCTAAACTATGAGGTCTAGCTATAA 180

Db 174 AAAGTCTCGGAGAAATCGTCTATGTATGAAGAGAAAGTATGAGGCCATGACTATAA 233

Qy 181 CTAGTTTCAAGGTACCTCCACCTTTCATGCTAGTAAACGGCTGAGACTTCCAC 240

Db 234 CTAGTTTCAAGGCTCCCTCCATCTTTCATGCTAAATAACGGGTCACAGACTTCCAG 293

Qy 241 GGGAAATGATTTTGGTAAACGATCGAAACACACAGGAATCAGGTTGAACCTCTCAGATGACT 300

Db 294 GGGAAATGATTTTGGTAAACGATCGAAACACACAGGAATCAGGTTGAACCTCTCAGATGACT 353

Qy 301 TTCGGCAGCTCCAGAGAAATCTTCCGGAAGATCATGCCCAAGAACGACGAGAGAGAA 360

Db 354 TTCGGCAGCTCCAGAGAAATCTTCCGGAAGATCATGCCCAAGAACGACGAGAGAGAA 413

Qy 361 AATGTTTGAAGGAAGTCCAGAGGATCTGGCCCAACAAAATGATGGAAACAGCTGTGC 420

Db 414 AATGTTTGAAGGAAGTCCAGAGGATCTGGCCCAACAAAATGATGGAAACAGCTGTGC 473

Qy 421 CCCCCGGGAATCCAAAGTACCTTGGAGAAATTAACAAGACATCTGACCCCAAAAGGGGG 480

Db 474 CCCCCGGGAATCCAAAGTACCTTGGAGAAATTAACAAGACATCTGACCCCAAAAGGGGG 533

Qy 481 AAACATGCTGACCCACAGACTGCGTGAGAGAAAGCAGCTGG--TGGTTTATGAAGAGA 538

Db 534 GAAATGCTGACCCACAGACTGCGTGAGAGAAAGCAGCTGGTGGTGGTTTATGAAGAGA 593

Qy 539 TCAGCGACCTTGAGGAAGATGACGAGTAATCTCCCTCG 576

Db 594 TCAGCGACCTTGAGGAAGATGACGAGTAATCTCCCTCG 631

Search completed: March 31, 2004, 15:40:56

Job time : 441 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 31, 2004, 15:19:29 ; Search time 97 Seconds
(without alignments)
3295.378 Million cell updates/sec

Title: US-09-975-856-1

Perfect score: 576

Sequence: 1 ATGAACGGAGACAGCGCTT.....ATGACGAGTAATCCCTCG 576

Scoring table: IDENTITY NUC

Gapop 10.0 , Capext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:*

- 1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
- 2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
- 3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
- 4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
- 5: /cgn2_6/ptodata/2/ina/PTCUS_COMB.seq.*
- 6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	576	100.0	576	3	US-09-105-839D-5
2	576	100.0	576	3	US-08-851-138C-1
3	576	100.0	576	4	US-09-796-780-1
4	576	100.0	576	4	US-09-344-040C-5
5	576	100.0	576	4	US-09-833-039A-5
6	576	100.0	576	4	US-09-392-714-17
7	496.6	86.2	766	4	US-09-392-714-16
8	496.6	86.2	931	1	US-08-479-328-2
9	496.6	86.2	931	1	US-08-761-119-2
10	496.6	86.2	931	2	US-08-668-128B-2
11	496.6	86.2	931	2	US-08-505-445-2
12	496.6	86.2	931	3	US-08-959-625-2
13	496.6	86.2	931	3	US-09-008-466-2
14	496.6	86.2	931	3	US-08-580-980A-2
15	496.6	86.2	931	3	US-09-053-453-2
16	496.6	86.2	931	3	US-08-844-116A-2
17	496.6	86.2	931	3	US-09-105-839D-2
18	496.6	86.2	931	4	US-09-437-858-2
19	496.6	86.2	931	4	US-09-344-040C-2
20	496.6	86.2	931	4	US-09-833-039A-2
21	482.2	83.7	576	3	US-09-105-839D-6
22	482.2	83.7	576	3	US-08-851-138C-2
23	482.2	83.7	576	4	US-09-796-780-2
24	482.2	83.7	576	4	US-09-344-040C-6
25	482.2	83.7	576	4	US-09-833-039A-6
26	480	83.3	766	3	US-09-105-839D-1
27	480	83.3	766	4	US-09-344-040C-1

c	28	480	83.3	766	4	US-09-833-039A-1	Sequence 1, Appli
	29	50.2	8.7	7218	1	US-08-232-463-14	Sequence 14, Appli
	30	44.2	7.7	448	4	US-09-621-976-9992	Sequence 9992, Ap
	31	39.4	6.8	1097	4	US-09-976-594-115	Sequence 115, App
	32	35.4	6.1	478	4	US-09-621-976-3458	Sequence 3458, Ap
	33	34.6	6.0	2628	4	US-09-294-531B-5	Sequence 5, Appli
	34	33.8	5.9	1185	2	US-08-179-557-14	Sequence 14, Appli
	35	33.4	5.8	5761	4	US-09-620-312D-147	Sequence 147, App
	36	33.4	5.6	6107	4	US-09-976-594-557	Sequence 557, App
	37	32.4	5.6	484	4	US-09-621-976-1092	Sequence 1092, Ap
	38	32.4	5.6	1001	4	US-09-641-638-257	Sequence 257, App
	39	32.4	5.6	2143	4	US-08-861-578A-3	Sequence 3, Appli
	40	32.2	5.6	358	2	US-08-692-787-5	Sequence 5, Appli
	41	32.2	5.6	358	3	US-09-097-199-5	Sequence 5, Appli
	42	31.8	5.5	832	4	US-09-621-976-2813	Sequence 2813, Ap
	43	31.6	5.5	505	4	US-09-621-976-15639	Sequence 15639, A
	44	31.6	5.5	1686	4	US-09-489-039A-3471	Sequence 3471, Ap
	45	31.6	5.5	1965	4	US-09-489-039A-3333	Sequence 3333, Ap

ALIGNMENTS

RESULT 1

US-09-105-839D-5
; Sequence 5, Application US/09105839D
; Patent No. 6287756
; GENERAL INFORMATION:
; APPLICANT: Tureci, Ozlem
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Sahin, Ugur
; APPLICANT: Gure, Ali
; APPLICANT: Old, Lloyd J
; APPLICANT: Pfeunderschuh, Michael
; TITLE OF INVENTION: Method for Determining Presence of Cancer In A Sample By Deter
; FILE REFERENCE: LUD 5556
; CURRENT APPLICATION NUMBER: US/09/105,839D
; CURRENT FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: US 08/851,130
; PRIOR FILING DATE: 1997-05-05
; NUMBER OF SEQ ID NOS: 72
; SEQ ID NO 5
; LENGTH: 576
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-105-839D-5

Query Match 100.0%; Score 576; DB 3; Length 576;
Best Local Similarity 100.0%; Pred. No. 1.7e-184;
Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGAACGGAGACAGCGCTTTGCAAGGAGACCCAGGATGATGCTCAATATCAGAGAAG	60
DB	1	ATGAACGGAGACAGCGCTTTGCAAGGAGACCCAGGATGATGCTCAATATCAGAGAAG	60
QY	61	TTACGAAAGGCGCTTCGATGATATTCGCAATATCTTCTTAAGAAAGAGTGGGAAAGATG	120
DB	61	TTACGAAAGGCGCTTCGATGATATTCGCAATATCTTCTTAAGAAAGAGTGGGAAAGATG	120
QY	121	AAATCCCGGAGAAATCGTCTATGTATATGAGCTAACTATGAGGTGATGACTGACTAAA	180
DB	121	AAATCCCGGAGAAATCGTCTATGTATATGAGCTAACTATGAGGTGATGACTGACTAAA	180
QY	181	CTAGGTTTCAAGGTCACCCCTCCACCTTCATGCGTAGTAACACCGGCTGCAGACTTCCAC	240
DB	181	CTAGGTTTCAAGGTCACCCCTCCACCTTCATGCGTAGTAACACCGGCTGCAGACTTCCAC	240
QY	241	GGGAATGATTTTGGTAACGATCGAAACACACGGAATCAGGTGTAACGCTTCAGATGACT	300
DB	241	GGGAATGATTTTGGTAACGATCGAAACACACGGAATCAGGTGTAACGCTTCAGATGACT	300
QY	301	TTCCGACGCTCCAGAGAATCTTCCCGAAGATCATGCCAAGAACCCAGCAGAGGAAGAA	360

Db 301 TTCCGACGCTCCAGAGAAATCTTCCGAGATCATGCCAAGAGCCAGCAGAGGAAGAA 360
 Qy 361 AATGGTTTGAAGAGAGTGCAGAGGATCTGGCCCAAAATGATGGAAACAGCTGTGC 420
 Db 361 AATGGTTTGAAGAGAGTGCAGAGGATCTGGCCCAAAATGATGGAAACAGCTGTGC 420
 Qy 421 CCCCCGGGAAATCCAAAGTACCTTGGAGAAATTAACAAAGACATCTGGACCCAAAGGGG 480
 Db 421 CCCCCGGGAAATCCAAAGTACCTTGGAGAAATTAACAAAGACATCTGGACCCAAAGGGG 480
 Qy 481 AAACATGCTGGACCCCAAGACTCGGTGAGAGAAAGCAGCTGGTGTATGAAGAGATC 540
 Db 481 AAACATGCTGGACCCCAAGACTCGGTGAGAGAAAGCAGCTGGTGTATGAAGAGATC 540
 Qy 541 AGCAGCCCTGAGGAAGATGACGAGTAATCTCCCTCG 576
 Db 541 AGCAGCCCTGAGGAAGATGACGAGTAATCTCCCTCG 576

RESULT 2

US-08-851-138C-1
 ; Sequence 1, Application US/08851138C
 ; Patent No. 6291658
 ; GENERAL INFORMATION:
 ; APPLICANT: Gure, Ali O.; Tureci, Ozlem; Sahin, Ugur; Tsang, Solam;
 ; APPLICANT: Scanlan, Matthew J.; Knuth Alexander, Pfreundschuh, Michael; Old,
 ; APPLICANT: Lloyd J.; Chen, Yao-Tseng
 ; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding SSX Family
 ; TITLE OF INVENTION: Members And Uses Thereof
 ; NUMBER OF SEQUENCES: 14
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Fulbright & Jaworski L.L.P.
 ; STREET: 666 Fifth Avenue
 ; CITY: New York City
 ; STATE: New York
 ; ZIP: 10103
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
 ; COMPUTER: IBM
 ; OPERATING SYSTEM: PC-DOS
 ; SOFTWARE: Wordperfect
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/851,138C
 ; FILING DATE: 5-May-1997
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Hanson, No. 6291658man D.
 ; REGISTRATION NUMBER: 30,946
 ; REFERENCE/DOCKET NUMBER: LUD 5480
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 318-3168
 ; TELEFAX: (212) 752-5958
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 576 nucleotides
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; US-08-851-138C-1

Query Match 100.0%; Score 576; DB 3; Length 576;
 Best Local Similarity 100.0%; Pred. No. 1.7e-184;
 Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 ATGAACGAGACGCGCTTTGCAAGGAGACCCAGGATGATCTCAATATCAGAGAAG 60
 Db 1 ATGAACGAGACGCGCTTTGCAAGGAGACCCAGGATGATCTCAATATCAGAGAAG 60
 Qy 61 TTACGAAGGCTTCGATGATATTGCCAATATCTCTAAGAAAGATGGGAAGATG 120
 Db 61 TTACGAAGGCTTCGATGATATTGCCAATATCTCTAAGAAAGATGGGAAGATG 120

Qy 121 AAATCCTCGGAGAAATCGTCTATGTGTATATGAAGCTAAACTATAGGTCATCACTAAA 180
 Db 121 AAATCCTCGGAGAAATCGTCTATGTGTATATGAAGCTAAACTATAGGTCATCACTAAA 180
 Qy 181 CTAGGTTTCAAGGTCACCCCTCCACCTTTCATGCTGTAGTAAACCGGCTGCAGACTTCAC 240
 Db 181 CTAGGTTTCAAGGTCACCCCTCCACCTTTCATGCTGTAGTAAACCGGCTGCAGACTTCAC 240
 Qy 241 GGGATGATTTTGGTTAAACGATCGAAACCCACAGGAATCAGGTTTGAACGCTCTCAGATGACT 300
 Db 241 GGGATGATTTTGGTTAAACGATCGAAACCCACAGGAATCAGGTTTGAACGCTCTCAGATGACT 300
 Qy 301 TTCGCGACCTCCAGAGATCTTCCGAGAGATCTGCCAAGAGATGATGAGAGAGAGAGAA 360
 Db 301 TTCGCGACCTCCAGAGATCTTCCGAGAGATCTGCCAAGAGATGATGAGAGAGAGAGAA 360
 Qy 361 AATGTTTGAAGGAAGTGCAGAGAGCATCTGCCCAACAAATGATGGGAAACAGCTGTGC 420
 Db 361 AATGTTTGAAGGAAGTGCAGAGAGCATCTGCCCAACAAATGATGGGAAACAGCTGTGC 420
 Qy 421 CCCCCGGGAAATCCAAAGTACCTTGGAGAAATTAACAAAGACATCTGGACCCAAAGGGG 480
 Db 421 CCCCCGGGAAATCCAAAGTACCTTGGAGAAATTAACAAAGACATCTGGACCCAAAGGGG 480
 Qy 481 AAACATGCTGGACCCCAAGACTCGGTGAGAGAAAGCAGCTGGTGTATGAAGAGATC 540
 Db 481 AAACATGCTGGACCCCAAGACTCGGTGAGAGAAAGCAGCTGGTGTATGAAGAGATC 540
 Qy 541 AGCAGCCCTGAGGAAGATGACGAGTAATCTCCCTCG 576
 Db 541 AGCAGCCCTGAGGAAGATGACGAGTAATCTCCCTCG 576

RESULT 3

US-09-796-780-1
 ; Sequence 1, Application US/09796780
 ; Patent No. 6339140
 ; GENERAL INFORMATION:
 ; APPLICANT: Gure, Ali O.; Tureci, Ozlem; Sahin, Ugur; Tsang, Solam;
 ; APPLICANT: Scanlan, Matthew J.; Knuth Alexander, Pfreundschuh, Michael; Old,
 ; APPLICANT: Lloyd J.; Chen, Yao-Tseng
 ; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding SSX
 ; TITLE OF INVENTION: Family
 ; TITLE OF INVENTION: Members And Uses Thereof
 ; NUMBER OF SEQUENCES: 14
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Fulbright & Jaworski L.L.P.
 ; STREET: 666 Fifth Avenue
 ; CITY: New York City
 ; STATE: New York
 ; ZIP: 10103
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
 ; COMPUTER: IBM
 ; OPERATING SYSTEM: PC-DOS
 ; SOFTWARE: Wordperfect
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/796,780
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/851,138
 ; FILING DATE: 5-May-1997
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Hanson, No. 6339140man D.
 ; REGISTRATION NUMBER: 30,946
 ; REFERENCE/DOCKET NUMBER: LUD 5480
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 318-3168
 ; TELEFAX: (212) 752-5958
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 576 nucleotides

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-796-780-1

Query Match 100.0%; Score 576; DB 4; Length 576;
Best Local Similarity 100.0%; Pred. No. 1.7e-184;
Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAACGGAGACGACGCCCTTTGCAAGGAGACCCAGGGATGATGCTCAAAATATCAGAGAAG 60
DB 1 ATGAACGGAGACGACGCCCTTTGCAAGGAGACCCAGGGATGATGCTCAAAATATCAGAGAAG 60
QY 61 TTACGAAAGGCCCTTCGATGATATGCAAAATATCTTCTTAAGAAAGAGTGGGAAAGATG 120
DB 61 TTACGAAAGGCCCTTCGATGATATGCAAAATATCTTCTTAAGAAAGAGTGGGAAAGATG 120
QY 121 AAATCTCGGAGAAATCGTCTATGTATATGAAGCTAACTATGAGGTCAATGACTAAA 180
DB 121 AAATCTCGGAGAAATCGTCTATGTATATGAAGCTAACTATGAGGTCAATGACTAAA 180
QY 181 CTAGGTTTCAAGGTCAACCTCCACCTTTTCATCGGTAGTAAACGGGCTGCAGACTTCCAC 240
DB 181 CTAGGTTTCAAGGTCAACCTCCACCTTTTCATCGGTAGTAAACGGGCTGCAGACTTCCAC 240
QY 241 GGGATGATTTTGGTAAAGTCAAGGATCGAAACACAGGATCAGGTTGAAGCTCCTCAGATGACT 300
DB 241 GGGATGATTTTGGTAAAGTCAAGGATCGAAACACAGGATCAGGTTGAAGCTCCTCAGATGACT 300
QY 301 TTCCGAGCCTCCAGAGATCTTCCGAGAGATCATGCCAAGAACGACAGAGAGAA 360
DB 301 TTCCGAGCCTCCAGAGATCTTCCGAGAGATCATGCCAAGAACGACAGAGAGAA 360
QY 361 AATGTTTGAAGAGTGCAGAGGATCTGCGCCACAAATATGAGGAAACAGCTGTGC 420
DB 361 AATGTTTGAAGAGTGCAGAGGATCTGCGCCACAAATATGAGGAAACAGCTGTGC 420
QY 421 CCCCCGGGAAATCCAGTACCTTGGAGAGATTAAACAGACATCTGGACCCAAAAGGGG 480
DB 421 CCCCCGGGAAATCCAGTACCTTGGAGAGATTAAACAGACATCTGGACCCAAAAGGGG 480
QY 481 AAACATGCTGACCCACAGATGCGTGAGAGAAAGCAGCTGGTGTATGAAGATC 540
DB 481 AAACATGCTGACCCACAGATGCGTGAGAGAAAGCAGCTGGTGTATGAAGATC 540
QY 541 AGCGACCTCGAGAGATGACGAGTAACCTCCCTCG 576
DB 541 AGCGACCTCGAGAGATGACGAGTAACCTCCCTCG 576

RESULT 4
US-09-344-040C-5
; Sequence 5, Application US/09344040C
; Patent No. 6548064
; GENERAL INFORMATION:
; APPLICANT: Tureci, Ozlem
; APPLICANT: Sahin, Ugur
; APPLICANT: Pfreundschuh, Michael
; APPLICANT: Ramensschuh, Hans Georg
; APPLICANT: Stevanovic, Stefan
; TITLE OF INVENTION: Method for Determining Presence of Cancer in a Sample By Determin
; TITLE OF INVENTION: Expression of an SSX Gene, Peptides Derived From Said SSX Gene a
; TITLE OF INVENTION: Gene, and Uses Thereof
; FILE REFERENCE: LUD 5556.1
; CURRENT FILING DATE: 1999-06-25
; PRIOR FILING DATE: 1999-06-25
; PRIOR FILING DATE: 1999-06-26
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: US 09/105,839
; PRIOR APPLICATION NUMBER: US 08/851,130
; NUMBER OF SEQ ID NOS: 132
; SEQ ID NO 5
; LENGTH: 576
; TYPE: DNA
; ORGANISM: Homo sapiens

TYPE: DNA
ORGANISM: Homo sapiens
US-09-344-040C-5

Query Match 100.0%; Score 576; DB 4; Length 576;
Best Local Similarity 100.0%; Pred. No. 1.7e-184;
Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAACGGAGACGACGCCCTTTGCAAGGAGACCCAGGGATGATGCTCAAAATATCAGAGAAG 60
DB 1 ATGAACGGAGACGACGCCCTTTGCAAGGAGACCCAGGGATGATGCTCAAAATATCAGAGAAG 60
QY 61 TTACGAAAGGCCCTTCGATGATATGCAAAATATCTTCTTAAGAAAGAGTGGGAAAGATG 120
DB 61 TTACGAAAGGCCCTTCGATGATATGCAAAATATCTTCTTAAGAAAGAGTGGGAAAGATG 120
QY 121 AAATCTCGGAGAAATCGTCTATGTATATGAAGCTAACTATGAGGTCAATGACTAAA 180
DB 121 AAATCTCGGAGAAATCGTCTATGTATATGAAGCTAACTATGAGGTCAATGACTAAA 180
QY 181 CTAGGTTTCAAGGTCAACCTCCACCTTTTCATCGGTAGTAAACGGGCTGCAGACTTCCAC 240
DB 181 CTAGGTTTCAAGGTCAACCTCCACCTTTTCATCGGTAGTAAACGGGCTGCAGACTTCCAC 240
QY 241 GGGATGATTTTGGTAAAGTCAAGGATCGAAACACAGGATCAGGTTGAAGCTCCTCAGATGACT 300
DB 241 GGGATGATTTTGGTAAAGTCAAGGATCGAAACACAGGATCAGGTTGAAGCTCCTCAGATGACT 300
QY 301 TTCCGAGCCTCCAGAGATCTTCCGAGAGATCATGCCAAGAACGACAGAGAGAA 360
DB 301 TTCCGAGCCTCCAGAGATCTTCCGAGAGATCATGCCAAGAACGACAGAGAGAA 360
QY 361 AATGTTTGAAGAGTGCAGAGGATCTGCGCCACAAATATGAGGAAACAGCTGTGC 420
DB 361 AATGTTTGAAGAGTGCAGAGGATCTGCGCCACAAATATGAGGAAACAGCTGTGC 420
QY 421 CCCCCGGGAAATCCAGTACCTTGGAGAGATTAAACAGACATCTGGACCCAAAAGGGG 480
DB 421 CCCCCGGGAAATCCAGTACCTTGGAGAGATTAAACAGACATCTGGACCCAAAAGGGG 480
QY 481 AAACATGCTGACCCACAGATGCGTGAGAGAAAGCAGCTGGTGTATGAAGATC 540
DB 481 AAACATGCTGACCCACAGATGCGTGAGAGAAAGCAGCTGGTGTATGAAGATC 540
QY 541 AGCGACCTCGAGAGATGACGAGTAACCTCCCTCG 576
DB 541 AGCGACCTCGAGAGATGACGAGTAACCTCCCTCG 576

RESULT 5
US-09-833-039A-5
; Sequence 5, Application US/09833039A
; Patent No. 6673350
; GENERAL INFORMATION:
; APPLICANT: Tureci, Ugur
; APPLICANT: Sahin, Ugur
; APPLICANT: Pfreundschuh, Michael
; TITLE OF INVENTION: Tumor Associated Peptide and Uses Thereof
; FILE REFERENCE: LUD 5622.1
; CURRENT APPLICATION NUMBER: US/09/833,039A
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: US 09/409,455
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: US 09/344,040
; PRIOR FILING DATE: 1999-06-26
; PRIOR APPLICATION NUMBER: US 09/105,839
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 5
; LENGTH: 576
; TYPE: DNA
; ORGANISM: Homo sapiens

US-09-833-039A-5

Query Match
Best Local Similarity 100.0%; Score 576; DB 4; Length 576;
Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAACGGAGACGACGCTTTGCAAGAGACCCAGGAGTATGCTCAAAATATCAGAGAAG 60
DB 1 ATGAACGGAGACGACGCTTTGCAAGAGACCCAGGAGTATGCTCAAAATATCAGAGAAG 60

QY 61 TTACGAAGGCTTCGATGATATTGCAAAATCTTCTTAAGAAAGTGGGAAAGATG 120
DB 61 TTACGAAGGCTTCGATGATATTGCAAAATCTTCTTAAGAAAGTGGGAAAGATG 120

QY 121 AAATCTCGAGAAAATCGTCTATGTATATGAAGCTAAACTATGAGTCTATGACTAAA 180
DB 121 AAATCTCGAGAAAATCGTCTATGTATATGAAGCTAAACTATGAGTCTATGACTAAA 180

QY 181 CTAGGTTTCAAGTCAACCTCCACCTTTTCATGCGTAGTAAACGGGCTGCAGACTTCCAC 240
DB 181 CTAGGTTTCAAGTCAACCTCCACCTTTTCATGCGTAGTAAACGGGCTGCAGACTTCCAC 240

QY 241 GGAATGATTTTGGTAAAGTCAAGATCGAAGTTCAGGTTGAAGCTCTCAGATGACT 300
DB 241 GGAATGATTTTGGTAAAGTCAAGATCGAAGTTCAGGTTGAAGCTCTCAGATGACT 300

QY 301 TTCCGAGGCTCCAGAGATCTTCCGAAAGTATGCGGAGATCATGCCCAAGAGCAAGCAAGAGAA 360
DB 301 TTCCGAGGCTCCAGAGATCTTCCGAAAGTATGCGGAGATCATGCCCAAGAGCAAGCAAGAGAA 360

QY 361 AATGGTTTGAAGGAGTGCAGAGGATCTGCGGAGAGATTAACAGACATCTGGACCCAAAGGGGG 420
DB 361 AATGGTTTGAAGGAGTGCAGAGGATCTGCGGAGAGATTAACAGACATCTGGACCCAAAGGGGG 420

QY 421 CCCCCGGGAAATCCAAAGTACCTTGGAGAGATTAACAGACATCTGGACCCAAAGGGGG 480
DB 421 CCCCCGGGAAATCCAAAGTACCTTGGAGAGATTAACAGACATCTGGACCCAAAGGGGG 480

QY 481 AAACATGCTGGACCCACAGATCTGCGGAGAGAGTAAAGAGATCTGGAGAGAGATC 540
DB 481 AAACATGCTGGACCCACAGATCTGCGGAGAGAGTAAAGAGATCTGGAGAGAGATC 540

QY 541 AGCGACCTCGAGAGATGACGAGTAACTCCCTCTCG 576
DB 541 AGCGACCTCGAGAGATGACGAGTAACTCCCTCTCG 576

RESULT 6

US-09-392-714-17
; Sequence 17, Application US/09392714A
; Patent No. 6686147
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew J.
; APPLICANT: Gure, Ali O.
; APPLICANT: Williamson, Barbara
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Cancer Associated Antigens and Uses
; FILE REFERENCE: 10461/7062
; CURRENT APPLICATION NUMBER: US/09/392,714A
; EARLIER FILING DATE: 1999-09-09
; EARLIER FILING DATE: PCT/US98/14679
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 17
; LENGTH: 576
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-392-714-17

Query Match 100.0%; Score 576; DB 4; Length 576;

Best Local Similarity 100.0%; Pred. No. 1.7e-184;
Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAACGGAGACGACGCTTTGCAAGAGACCCAGGAGTATGCTCAAAATATCAGAGAAG 60
DB 1 ATGAACGGAGACGACGCTTTGCAAGAGACCCAGGAGTATGCTCAAAATATCAGAGAAG 60

QY 61 TTACGAAGGCTTCGATGATATTGCAAAATCTTCTTAAGAAAGTGGGAAAGATG 120
DB 61 TTACGAAGGCTTCGATGATATTGCAAAATCTTCTTAAGAAAGTGGGAAAGATG 120

QY 121 AAATCTCGGAGAAAATCGTCTATGTATATGAAGCTAAACTATGAGTCTATGACTAAA 180
DB 121 AAATCTCGGAGAAAATCGTCTATGTATATGAAGCTAAACTATGAGTCTATGACTAAA 180

QY 181 CTAGGTTTCAAGTCAACCTCCACCTTTTCATGCGTAGTAAACGGGCTGCAGACTTCCAC 240
DB 181 CTAGGTTTCAAGTCAACCTCCACCTTTTCATGCGTAGTAAACGGGCTGCAGACTTCCAC 240

QY 241 GGAATGATTTTGGTAAAGTCAAGATCGAAGTTCAGGTTGAAGCTCTCAGATGACT 300
DB 241 GGAATGATTTTGGTAAAGTCAAGATCGAAGTTCAGGTTGAAGCTCTCAGATGACT 300

QY 301 TTCCGAGGCTCCAGAGATCTTCCGAAAGTATGCGGAGATCATGCCCAAGAGCAAGCAAGAGAA 360
DB 301 TTCCGAGGCTCCAGAGATCTTCCGAAAGTATGCGGAGATCATGCCCAAGAGCAAGCAAGAGAA 360

QY 361 AATGGTTTGAAGGAGTGCAGAGGATCTGCGGAGAGATTAACAGACATCTGGACCCAAAGGGGG 420
DB 361 AATGGTTTGAAGGAGTGCAGAGGATCTGCGGAGAGATTAACAGACATCTGGACCCAAAGGGGG 420

QY 421 CCCCCGGGAAATCCAAAGTACCTTGGAGAGATTAACAGACATCTGGACCCAAAGGGGG 480
DB 421 CCCCCGGGAAATCCAAAGTACCTTGGAGAGATTAACAGACATCTGGACCCAAAGGGGG 480

QY 481 AAACATGCTGGACCCACAGATCTGCGGAGAGAGTAAAGAGATCTGGAGAGAGATC 540
DB 481 AAACATGCTGGACCCACAGATCTGCGGAGAGAGTAAAGAGATCTGGAGAGAGATC 540

QY 541 AGCGACCTCGAGAGATGACGAGTAACTCCCTCTCG 576
DB 541 AGCGACCTCGAGAGATGACGAGTAACTCCCTCTCG 576

RESULT 7

US-09-392-714-16
; Sequence 16, Application US/09392714A
; Patent No. 6686147
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew J.
; APPLICANT: Gure, Ali O.
; APPLICANT: Williamson, Barbara
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Cancer Associated Antigens and Uses
; FILE REFERENCE: 10461/7062
; CURRENT APPLICATION NUMBER: US/09/392,714A
; EARLIER FILING DATE: 1999-09-09
; EARLIER FILING DATE: PCT/US98/14679
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 16
; LENGTH: 766
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-392-714-16

Query Match 86.2%; Score 496.6; DB 4; Length 766;
Best Local Similarity 91.5%; Pred. No. 1.3e-157;
Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 1 ATGAACGGAGACGACGCCCTTTCAGGAGACCCAGGATGATGCTCAATATCAGAGAAG 60
Db 92 ATGAACGGAGACGACGCCCTTTCAGGAGACCCAGGATGATGCTCAATATCAGAGAAG 151
QY 61 TTACGAAGAGCCTTCGATGATATTCGCAAAATCTCTTAAGAAGAGTGGGAAAGATG 120
Db 152 ATCCAAAAGGCCTTCGATGATATTCGCAAAATCTCTTAAGAAGAGTGGGAAAGATG 211
QY 121 AAATCCTCGGAGAAATCGTCTATGTGTATGAAGCTAAACTATGAGGTCTACACTAAA 180
Db 212 AAAGCCTCGGAGAAATCTCTATGTGTATGAAGAGAAAGTATGAGGCTATGCTAAA 271
QY 181 CTAGTTCAGAGTCCACCTCCACCTTTTCATGTGTATGAAGCTAAACTATGAGGTCTACACTAAA 240
Db 272 CTAGTTCAGAGTCCACCTCCACCTTTTCATGTGTATGAAGAGAAAGTATGAGGCTATGCTAAA 331
QY 241 GGGATGATTTTGGTAAAGATCGAATCCGAAACACAGGATCAGGTTGAACGTCCTCAGATGACT 300
Db 332 GGGATGATTTTGGTAAAGATCGAATCCGAAACACAGGATCAGGTTGAACGTCCTCAGATGACT 391
QY 301 TTCCGAGCCTCCAGAGATCTTCCGAGAGATCATGCCCAAGAGCCAGCAGAGAGAGAA 360
Db 392 TTCCGAGCCTCCAGAGATCTTCCGAGAGATCATGCCCAAGAGCCAGCAGAGAGAGAA 451
QY 361 AATGCTTTGAAGAGTCCAGAGCATCTGCCCAAGATCATGCCCAAGAGCCAGCAGAGAGAG 420
Db 452 AATGCTTTGAAGAGTCCAGAGCATCTGCCCAAGATCATGCCCAAGAGCCAGCAGAGAGAG 511
QY 421 CCCCCGGGAAATCCAAAGTACCTTTGGAGAGATTAACAAGACATCTGGACCCCAAGAGGGG 480
Db 512 CCCCCGGGAAATCCAAAGTACCTTTGGAGAGATTAACAAGACATCTGGACCCCAAGAGGGG 571
QY 481 AAACATGCTGACCCAGACAGTCTGCTGAGAGAAAGCAGCTGGTGTATGAAGATC 540
Db 572 GAACATGCTGACCCAGACAGTCTGCTGAGAGAAAGCAGCTGGTGTATGAAGATC 631
QY 541 AGCGACCTCGAGAGAGTACGAGTAACCTCCCTC 575
Db 632 AGCGACCTCGAGAGAGTACGAGTAACCTCCCTC 666

RESULT 8

US-08-479-328-2
; Sequence 2, Application US/08479328
; Patent No. 5698396
; GENERAL INFORMATION:
; APPLICANT: Pfreundschuh, Michael
; TITLE OF INVENTION: Method For Identifying Or Isolating A Molecule
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/479,328
; FILING DATE: 07-JUNE-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5698396man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5410
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 931 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double-stranded
; TOPOLOGY: linear
; US-08-479-328-2
Query Match 86.2%; Score 496.6; DB 1; Length 931;
Best Local Similarity 91.5%; Pred. No. 1.5e-157;
Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
QY 1 ATGAACGGAGACGACGCCCTTTCAGGAGACCCAGGATGATGCTCAATATCAGAGAAG 60
Db 99 ATGAACGGAGACGACGCCCTTTCAGGAGACCCAGGATGATGCTCAATATCAGAGAAG 158
QY 61 TTACGAAGAGCCTTCGATGATATTCGCAAAATCTCTTAAGAAGAGTGGGAAAGATG 120
Db 159 ATCCAAAAGGCCTTCGATGATATTCGCAAAATCTCTTAAGAAGAGTGGGAAAGATG 218
QY 121 AAATCCTCGGAGAAATCGTCTATGTGTATGAAGCTAAACTATGAGGTCTATGACTAAA 180
Db 219 AAAGCCTCGGAGAAATCTCTATGTGTATGAAGAGAAAGTATGAGGCTATGACTAAA 278
QY 181 CTAGTTCAGAGTCCACCTCCACCTTTTCATGCGTAGTAAACGGGCTCAGACTTCCAC 240
Db 279 CTAGTTCAGAGTCCACCTCCACCTTTTCATGCGTAGTAAACGGGCTCAGACTTCCAC 338
QY 241 GGGATGATTTTGGTAAAGATCGAATCCGAAACACAGGATCAGGTTGAACGTCCTCAGATGACT 300
Db 339 GGGATGATTTTGGTAAAGATCGAATCCGAAACACAGGATCAGGTTGAACGTCCTCAGATGACT 398
QY 301 TTCCGAGCCTCCAGAGATCTTCCGAGAGATCATGCCCAAGAGCCAGCAGAGAGAGAA 360
Db 399 TTCCGAGCCTCCAGAGATCTTCCGAGAGATCATGCCCAAGAGCCAGCAGAGAGAGAA 458
QY 361 AATGCTTTGAAGAGTCCAGAGCATCTGCCCAAGATTAACAAGACATCTGGACCCCAAGAGGGG 420
Db 459 AATGCTTTGAAGAGTCCAGAGCATCTGCCCAAGATTAACAAGACATCTGGACCCCAAGAGGGG 518
QY 421 CCCCCGGGAAATCCAAAGTACCTTTGGAGAGATTAACAAGACATCTGGACCCCAAGAGGGG 480
Db 519 CCCCCGGGAAATCCAAAGTACCTTTGGAGAGATTAACAAGACATCTGGACCCCAAGAGGGG 578
QY 481 AAACATGCTGACCCAGACAGTCTGCTGAGAGAAAGCAGCTGGTGTATGAAGATC 540
Db 579 GAACATGCTGACCCAGACAGTCTGCTGAGAGAAAGCAGCTGGTGTATGAAGATC 638
QY 541 AGCGACCTCGAGAGAGTACGAGTAACCTCCCTC 575
Db 639 AGCGACCTCGAGAGAGTACGAGTAACCTCCCTC 673

RESULT 9

US-08-761-119-2
; Sequence 2, Application US/08761119
; Patent No. 5798264
; GENERAL INFORMATION:
; APPLICANT: Pfreundschuh, Michael
; TITLE OF INVENTION: Method For Identifying Or Isolating A Molecule
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/761,119
 FILING DATE: 6-DECEMBER-1996
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/479,328
 FILING DATE: 07-JUNE-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Hanson, No. 5798264man D.
 REGISTRATION NUMBER: 30,946
 REFERENCE/DOCKET NUMBER: LUD 5410.3
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 688-9200
 TELEFAX: (212) 838-3884
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 931 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double-stranded
 TOPOLOGY: linear
 US-08-761-119-2

Query Match 86.2%; Score 496.6; DB 1; Length 931;
 Best Local Similarity 91.5%; Pred. No. 1.5e-157;
 Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY	1	ATGAACGGAGACGACGCTTTCGAGGACACCCAGGATGCTCAATATCAGAGAAG	60
Db	99	ATGAACGGAGACGACGCTTTCGAGGACACCCAGGATGCTCAATATCAGAGAAG	158
QY	61	TTACGAAAGGCTTCGATGATATTCCTCAATATCTCTTAAGAAAGAGTGGGAAAGATG	120
Db	159	ATCCAAAGGCTTCGATGATATTCCTCAATATCTCTTAAGAAAGAGTGGGAAAGATG	218
QY	121	AAATCTCGGAGAAATCGTCTATGTATATGAAGCTAACTATGAGTCTAGCTAA	180
Db	219	AAAGCTTCGGAGAAATCTCTATGTATATGAAGAAAGATGAGGCTATGACTAA	278
QY	181	CTAGGTTTCAAGTCAACCTCCACCTTTCATCGCTAGTAAACGGGCTGCAGACTCCAC	240
Db	279	CTAGGTTTCAAGGCTCCACCTCCACCTTTCATGTATATGAAGAAAGATGAGGCTCCAG	338
QY	241	GGGAATGATTGGTAATGATGAAACACAGAAATCAGTTTGAACGCTCAGATGACT	300
Db	339	GGGAATGATTGGTAATGATGAAACACAGAAATCAGTTTGAACGCTCAGATGACT	398
QY	301	TTCCGAGGCTCCAGGAAATCTCCGAGGATCATGCCCAAGAGTCCAGAGGAGAA	360
Db	399	TTCCGAGGCTCCAGGAAATCTCCGAGGATCATGCCCAAGAGTCCAGAGGAGAA	458
QY	361	AATGGTTTGAAGAAATGCGAGGATCTGGCCCAAAATATGATGGGAAACAGCTGTGC	420
Db	459	AATGGTTTGAAGAAATGCGAGGATCTGGCCCAAAATATGATGGGAAACAGCTGTGC	518
QY	421	CCCCGGGAAATCCAAAGTACCTTGGAGAGATTAACAAGACATCTGGACCCAAAGGGG	480
Db	519	CCCCGGGAAATCCAAAGTACCTTGGAGAGATTAACAAGACATCTGGACCCAAAGGGG	578
QY	481	AAACATGCTGGACCCACAGATCTGCTGAGAAAGACAGCTGGTGGTTATGAAGATC	540
Db	579	GAACATGCTGGACCCACAGATCTGCTGAGAAAGACAGCTGGTGGTTATGAAGATC	638
QY	541	AGCGACCTTCAGGAAGATCAGGATTAACCTCCCTC	575
Db	639	AGCGACCTTCAGGAAGATCAGGATTAACCTCCCTC	673

RESULT 10
 US-08-668-128B-2
 Sequence 2, Application US/08668128B
 Patent No. 5840568
 GENERAL INFORMATION:
 APPLICANT: Pfeunderschuh, Michael
 TITLE OF INVENTION: Hodgkin's Disease Associated Molecules And

TITLE OF INVENTION: Uses Thereof
 NUMBER OF SEQUENCES: 14
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Felfe & Lynch
 STREET: 805 Third Avenue
 CITY: New York City
 STATE: New York
 ZIP: 10022
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
 COMPUTER: IBM
 OPERATING SYSTEM: PC-DOS
 SOFTWARE: Wordperfect
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/668,128B
 FILING DATE: 21-JUNE-1996
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/644,116
 FILING DATE: 10-MAY-1996
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/580,980
 FILING DATE: 03-JANUARY-1996
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/479,328
 FILING DATE: 07-JUNE-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Hanson, No. 5840568man D.
 REGISTRATION NUMBER: 30,946
 REFERENCE/DOCKET NUMBER: LUD 5441
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 688-9200
 TELEFAX: (212) 838-3884
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 931 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double-stranded
 TOPOLOGY: linear
 US-08-668-128B-2

Query Match 86.2%; Score 496.6; DB 2; Length 931;
 Best Local Similarity 91.5%; Pred. No. 1.5e-157;
 Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY	1	ATGAACGGAGACGACGCTTTCGAGGACACCCAGGATGATGCTCAATATCAGAGAAG	60
Db	99	ATGAACGGAGACGACGCTTTCGAGGACACCCAGGATGATGCTCAATATCAGAGAAG	158
QY	61	TTACGAAAGGCTTCGATGATATTCCTCAATATCTCTTAAGAAAGAGTGGGAAAGATG	120
Db	159	ATCCAAAGGCTTCGATGATATTCCTCAATATCTCTTAAGAAAGAGTGGGAAAGATG	218
QY	121	AAATCTCGGAGAAATCGTCTATGTATATGAAGCTAACTATGAGTCTAGCTAA	180
Db	219	AAAGCTTCGGAGAAATCTCTATGTATATGAAGAAAGATGAGGCTATGACTAA	278
QY	181	CTAGGTTTCAAGTCAACCTCCACCTTTCATCGCTAGTAAACGGGCTGCAGACTCCAC	240
Db	279	CTAGGTTTCAAGGCTCCACCTCCACCTTTCATGTATATGAAGAAAGATGAGGCTCCAG	338
QY	241	GGGAATGATTGGTAACGATCGAAACACAGAAATCAGTTTGAACGCTCAGATGACT	300
Db	339	GGGAATGATTGGTAACGATCGAAACACAGAAATCAGTTTGAACGCTCAGATGACT	398
QY	301	TTCCGAGGCTCCAGGAAATCTCCGAGGATCATGCCCAAGAGTCCAGAGGAGAA	360
Db	399	TTCCGAGGCTCCAGGAAATCTCCGAGGATCATGCCCAAGAGTCCAGAGGAGAA	458
QY	361	AATGGTTTGAAGAAATGCGAGGATCTGGCCCAAAATATGATGGGAAACAGCTGTGC	420
Db	459	AATGGTTTGAAGAAATGCGAGGATCTGGCCCAAAATATGATGGGAAACAGCTGTGC	518

QY 421 CCCCCGGGAATCCAAAGTACCTTGGAGAGATTAAACAGACATCTGGACCCCAAAAGGGG 480
DB 519 CCCCCGGGAAACCACTCTCTGAGAGATTACGAGAGATCTGGACCCCAAAAGGGG 578
QY 481 AAACATGCTGACCCACAGACTGCGTGAGAGAAAGCAGCTGGTGTTTATGAGAGAGATC 540
DB 579 GAACATGCTGACCCACAGACTGCGTGAGAGAAAGCAGCTGGTGTTTATGAGAGATC 638
QY 541 AGCGACCTTGAGGAAGATGACGAGTAACCTCCCTC 575
DB 639 AGCGACCTTGAGGAAGATGACGAGTAACCTCCCTC 673

RESULT 11

US-08-905-445-2
; Sequence 2, Application US/08905445
; Patent No. 5864015
; GENERAL INFORMATION:
; APPLICANT: Pfreundschuh, Michael
; TITLE OF INVENTION: Hodgkin's Disease Associated Molecules And
; TITLE OF INVENTION: Uses Thereof
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/905,445
; FILING DATE: 04-AUG-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/668,128
; FILING DATE: 21-JUNE-1996
; APPLICATION NUMBER: 08/644,116
; FILING DATE: 10-MAY-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/580,980
; FILING DATE: 03-JANUARY-1996
; APPLICATION NUMBER: 08/479,328
; FILING DATE: 07-JUNE-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5864015man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5441
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 931 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double-stranded
; TOPOLOGY: linear

US-08-905-445-2

Query Match 86.2%; Score 496.6; DB 2; Length 931;
Best Local Similarity 91.5%; Pred. No. 1.5e-157;
Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
QY 1 ATGACGGAGACGACGCTTGAAGGAGACCCAGGAGATGCTCAAAATATCAGAGAG 60
DB 99 ATGACGGAGACGACGCTTGAAGGAGACCCAGGAGATGCTCAAAATATCAGAGAG 158
QY 61 TTACGAAAGCCCTTCGATGATATTCGAAATCTCTTAAGAAAGATGGGAAAGATG 120

DB 159 ATCCAAAAGCCTTCGATGATATTGCCAAATACTTCTCTAAGAAAGATGGGAAAGATG 218
QY 121 AAATCTCGAGAGAAATCGTCTATGTATATGAAGCTAAACTATAGAGTCTATGACTAAA 180
DB 219 AAAGCCTCGAGAGAAATCTTCTATGTATATGAAGAGAAAGTATGAGGCTATGACTAAA 278
QY 181 CTAGGTTTCAAGGTCAACCTCCACCTTTCATGCGTAGTAAACGGGCTCGAGACTTCCAC 240
DB 279 CTAGGTTTCAAGGCAACCTCCACCTTTCATGTGTATTAACGGGCCGAGACTTCCAG 338
QY 241 GGAATGATTTTGGTAAACGATCGAAACACAGGAATCAGGTTGAACGCTCTCAGATGACT 300
DB 339 GGAATGATTTGATTAATGACCTTAACCGTGGGAATCAGGTTGAACGCTCTCAGATGACT 398
QY 301 TTCGGCAGCTCCAGAGAAATCTCCGGAAGATCATGCCCAAGAGCCACAGAGGAAGAA 360
DB 399 TTCGGCAGGCTCCAGGGAATCTCCCGAAGATCATGCCCAAGAGCCACAGAGGAGGA 458
QY 361 AATGGTTTGAAGGAAGTCCAGAGGATCTGGCCCAAAAATGATGGGAAACAGCTGTGC 420
DB 459 AATGATTCGGAGGAAGTCCAGAGGATCTGGCCCAAAAATGATGGGAAAGAGCTGTGC 518
QY 421 CCCCCGGGAAATCCAAAGTACCTTGGAGAGATTAAACAGACATCTGGACCCCAAAAGGGG 480
DB 519 CCCCCGGGAAACCACTCTCTGAGAGATTACGAGAGATCTGGACCCCAAAAGGGG 578
QY 481 AAACATGCTGACCCACAGACTGCGTGAGAGAAAGCAGCTGGTGTTTATGAGAGATC 540
DB 579 GAACATGCTGACCCACAGACTGCGTGAGAGAAAGCAGCTGGTGTTTATGAGAGATC 638
QY 541 AGCGACCTTGAGGAAGATGACGAGTAACCTCCCTC 575
DB 639 AGCGACCTTGAGGAAGATGACGAGTAACCTCCCTC 673

RESULT 12

US-08-959-625-2
; Sequence 2, Application US/08959625
; Patent No. 6017716
; GENERAL INFORMATION:
; APPLICANT: Pfreundschuh, Michael
; TITLE OF INVENTION: Method For Identifying Or Isolating A Molecule
; TITLE OF INVENTION: And Molecules Identified Thereby
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/959,625
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/580,980
; FILING DATE: 03-JANUARY-1996
; APPLICATION NUMBER: 08/479,328
; FILING DATE: 07-JUNE-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 6017716man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5410.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:

LENGTH: 931 base pairs
TYPE: nucleic acid
STRANDEDNESS: double-stranded
TOPOLOGY: linear
US-08-959-625-2

Query Match 86.2%; Score 496.6; DB 3; Length 931;
Best Local Similarity 91.5%; Pred. No. 1.5e-157;
Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 1 ATGAACGGAGACGACGCTTTCAGAGAGACCCAGGATGATGCTCAAAATATCAGAGAAG 60
Db 99 ATGAACGGAGACGACGCTTTCAGAGAGACCCAGGATGATGCTCAAAATATCAGAGAAG 158
QY 61 TTACGAAAGCCCTTCGATGATATGCGCAATATCTCTAAGAAAGAGTGGGAAAGATG 120
Db 159 ATCCAAAGGCCCTTCGATGATATGCGCAATATCTCTAAGAAAGAGTGGGAAAGATG 218
QY 121 AATCTCTCGAGAAATCGTCTATGTATATGAAGTAAATCTATGAGTCTAGTCTAA 180
Db 219 AAAGCCCTCGAGAAATCTCTATGTATATGAAGTAAATCTATGAGTCTAGTCTAA 278
QY 181 CTAGGTTTCAAGGTCACCTCCACACCTTTCATGCGTAGTAAACGGCTGCAGACTTCCAC 240
Db 279 CTAGGTTTCAAGGTCACCTCCACACCTTTCATGCGTAGTAAACGGCTGCAGACTTCCAC 338
QY 241 GGAATGATTTGGTAAAGATCGAAACCCAGGAATCAGGTTGAACGCTCTCAGATGACT 300
Db 339 GGAATGATTTGGTAAATATGACCTTAACCGTGGGAATCAGGTTGAACGCTCTCAGATGACT 398
QY 301 TTGGCAGCTCCAGAGAAATCTTCCGGAAGATCATGCCAAGAGCCAGCAGAGGAAGAA 360
Db 399 TTGGCAGCTCCAGAGAAATCTTCCGGAAGATCATGCCAAGAGCCAGCAGAGGAAGAA 458
QY 361 AATGGTTTGAAGGAATGCGCAGAGGATCTGGGACCAAAATATGATGGGAAACAGTGTGC 420
Db 459 AATGATTCGAGGAAGTGCAGAGGATCTGGGACCAAAATATGATGGGAAACAGTGTGC 518
QY 421 CCCCCGGGAATCCAGTACCTTGGAGAGATTAACAGACATCTGGACCCCAAGGGGG 480
Db 519 CCCCCGGGAATCCAGTACCTTGGAGAGATTAACAGACATCTGGACCCCAAGGGGG 578
QY 481 AAACATGCTGGACCCACAGACTGCGTGAAGAAAGCAGCTGGTGGTTTATGAAGATC 540
Db 579 GAACATGCTGGACCCACAGACTGCGTGAAGAAAGCAGCTGGTGGTTTATGAAGATC 638

RESULT 13

US-09-008-466-2
Sequence 2, Application US/09008466
Patent No. 6020134

GENERAL INFORMATION:

APPLICANT: Pfreundschuh, Michael
TITLE OF INVENTION: Method For Identifying Or Isolating A Molecule
TITLE OF INVENTION: And Molecules Identified Thereby

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: Felfe & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
ZIP: 10022

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
COMPUTER: IBM
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/008,466

QY 1 ATGAACGGAGACGACGCTTTCAGAGAGACCCAGGATGATGCTCAAAATATCAGAGAAG 60
Db 99 ATGAACGGAGACGACGCTTTCAGAGAGACCCAGGATGATGCTCAAAATATCAGAGAAG 158
QY 61 TTACGAAAGCCCTTCGATGATATGCGCAATATCTCTAAGAAAGAGTGGGAAAGATG 120
Db 159 ATCCAAAGGCCCTTCGATGATATGCGCAATATCTCTAAGAAAGAGTGGGAAAGATG 218
QY 121 AATCTCTCGAGAAATCGTCTATGTATATGAAGTAAATCTATGAGTCTAGTCTAA 180
Db 219 AAAGCCCTCGAGAAATCTCTATGTATATGAAGTAAATCTATGAGTCTAGTCTAA 278
QY 181 CTAGGTTTCAAGGTCACCTCCACACCTTTCATGCGTAGTAAACGGCTGCAGACTTCCAC 240
Db 279 CTAGGTTTCAAGGTCACCTCCACACCTTTCATGCGTAGTAAACGGCTGCAGACTTCCAC 338
QY 241 GGAATGATTTGGTAAAGATCGAAACCCAGGAATCAGGTTGAACGCTCTCAGATGACT 300
Db 339 GGAATGATTTGGTAAATATGACCTTAACCGTGGGAATCAGGTTGAACGCTCTCAGATGACT 398
QY 301 TTGGCAGCTCCAGAGAAATCTTCCGGAAGATCATGCCAAGAGCCAGCAGAGGAAGAA 360
Db 399 TTGGCAGCTCCAGAGAAATCTTCCGGAAGATCATGCCAAGAGCCAGCAGAGGAAGAA 458
QY 361 AATGGTTTGAAGGAATGCGCAGAGGATCTGGGACCAAAATATGATGGGAAACAGTGTGC 420
Db 459 AATGATTCGAGGAAGTGCAGAGGATCTGGGACCAAAATATGATGGGAAACAGTGTGC 518
QY 421 CCCCCGGGAATCCAGTACCTTGGAGAGATTAACAGACATCTGGACCCCAAGGGGG 480
Db 519 CCCCCGGGAATCCAGTACCTTGGAGAGATTAACAGACATCTGGACCCCAAGGGGG 578
QY 481 AAACATGCTGGACCCACAGACTGCGTGAAGAAAGCAGCTGGTGGTTTATGAAGATC 540
Db 579 GAACATGCTGGACCCACAGACTGCGTGAAGAAAGCAGCTGGTGGTTTATGAAGATC 638

Query Match

Best Local Similarity 86.2%; Score 496.6; DB 3; Length 931;
Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 1 ATGAACGGAGACGACGCTTTCAGAGAGACCCAGGATGATGCTCAAAATATCAGAGAAG 60
Db 99 ATGAACGGAGACGACGCTTTCAGAGAGACCCAGGATGATGCTCAAAATATCAGAGAAG 158
QY 61 TTACGAAAGCCCTTCGATGATATGCGCAATATCTCTAAGAAAGAGTGGGAAAGATG 120
Db 159 ATCCAAAGGCCCTTCGATGATATGCGCAATATCTCTAAGAAAGAGTGGGAAAGATG 218
QY 121 AATCTCTCGAGAAATCGTCTATGTATATGAAGTAAATCTATGAGTCTAGTCTAA 180
Db 219 AAAGCCCTCGAGAAATCTCTATGTATATGAAGTAAATCTATGAGTCTAGTCTAA 278
QY 181 CTAGGTTTCAAGGTCACCTCCACACCTTTCATGCGTAGTAAACGGCTGCAGACTTCCAC 240
Db 279 CTAGGTTTCAAGGTCACCTCCACACCTTTCATGCGTAGTAAACGGCTGCAGACTTCCAC 338
QY 241 GGAATGATTTGGTAAAGATCGAAACCCAGGAATCAGGTTGAACGCTCTCAGATGACT 300
Db 339 GGAATGATTTGGTAAATATGACCTTAACCGTGGGAATCAGGTTGAACGCTCTCAGATGACT 398
QY 301 TTGGCAGCTCCAGAGAAATCTTCCGGAAGATCATGCCAAGAGCCAGCAGAGGAAGAA 360
Db 399 TTGGCAGCTCCAGAGAAATCTTCCGGAAGATCATGCCAAGAGCCAGCAGAGGAAGAA 458
QY 361 AATGGTTTGAAGGAATGCGCAGAGGATCTGGGACCAAAATATGATGGGAAACAGTGTGC 420
Db 459 AATGATTCGAGGAAGTGCAGAGGATCTGGGACCAAAATATGATGGGAAACAGTGTGC 518
QY 421 CCCCCGGGAATCCAGTACCTTGGAGAGATTAACAGACATCTGGACCCCAAGGGGG 480
Db 519 CCCCCGGGAATCCAGTACCTTGGAGAGATTAACAGACATCTGGACCCCAAGGGGG 578
QY 481 AAACATGCTGGACCCACAGACTGCGTGAAGAAAGCAGCTGGTGGTTTATGAAGATC 540
Db 579 GAACATGCTGGACCCACAGACTGCGTGAAGAAAGCAGCTGGTGGTTTATGAAGATC 638
QY 541 AGCGACCTCGAGGAAGATGACGAGTAACTCCCTC 575
Db 639 AGCGACCTCGAGGAAGATGACGAGTAACTCCCTC 673

RESULT 14

US-08-580-980A-2
Sequence 2, Application US/08580980A
Patent No. 6025191
GENERAL INFORMATION:
APPLICANT: Pfreundschuh, Michael

339	GGGAATGATTTTGGATAATGACCTTAACCGTGGGAATCAGGTTGAAAGTCTCTCAGATGACT	398
301	TTTGGCAGGCTCCAGAGAATCTTCCGAAGATCATGCCAAGAACCCAGCAGAGGAAGA	360
399	TTTGGCAGGCTCCAGGGAATCTCCCGAAGATCATGCCAAGGAAGCCAGCAGAGGAAGA	458
361	AATGTTTGTGAAGGAAGTCCAGAGCATCTGGCCCAACAAAATGATGGGAAAACAGCTGTGC	420
459	AATGATTCGGAGGAAGTCCAGAAGCATCTGGCCCAACAAAATGATGGGAAGAGCTGTGC	518
421	CCCCCGGGAATCCAAGTACCTTCGAGAAGATTAAACAAGACATCTGGACCCCAAAAGGGGG	480
519	CCCCCGGGAACCAACTACTCTGAGNAGATTCAAGAGATCTGGAACCAAAAGGGGG	578
481	AAACATGCTTGACCCCAAGACTCGGTGAGNAGAAAGCAGCTGGTGTATTGAAAGATC	540
579	GAAATGCTTGACCCCAAGACTCGGTGAGNAGAAAACAGCTGGTGTATTGAAAGATC	638
541	AGGACCTTGAGGAAGATGACAGTAACCTCCCTC	575
639	AGCGACCTTGAGGAAGATGACAGTAACCTCCCTC	673

Search completed: March 31, 2004, 17:26:48
Job time : 99 secs


```
;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 576 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-975-856-1

Query Match      100.0%; Score 576; DB 10; Length 576;
Best Local Similarity 100.0%; Pred. No. 2.8e-174;
Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAACGGAGACGACGCTTTGCAAGAGACCCAGGGATGATGCTCAAAATATCAGAGAAG 60
DB 1 ATGAACGGAGACGACGCTTTGCAAGAGACCCAGGGATGATGCTCAAAATATCAGAGAAG 60
QY 61 TTACGAAAGGCCCTTCGATGATATGTCCTTAAGAAAGAGTGGGAAAGATG 120
DB 61 TTACGAAAGGCCCTTCGATGATATGTCCTTAAGAAAGAGTGGGAAAGATG 120
QY 121 AAATCCTCGAGAAATCGTCTATGTATATGAAGCTAACTATGAGTCAATGACTAAA 180
DB 121 AAATCCTCGAGAAATCGTCTATGTATATGAAGCTAACTATGAGTCAATGACTAAA 180
QY 181 CTAGGTTTCAAGGTCACCTCCACCTTTTCATGCTAGTAAACGGGCTGCAGACTTCCAC 240
DB 181 CTAGGTTTCAAGGTCACCTCCACCTTTTCATGCTAGTAAACGGGCTGCAGACTTCCAC 240
QY 241 GGGAAATGATTTGGTAAAGATCGAAGACCAAGGAATCAGGTTTCAAGCTTCTCAGATGACT 300
DB 241 GGGAAATGATTTGGTAAAGATCGAAGACCAAGGAATCAGGTTTCAAGCTTCTCAGATGACT 300
QY 301 TTCCGAGAGCTCCAGAGAAATCTTCCGAGAGATCATGCCCAAGAGCCAGCAGAGAGAA 360
DB 301 TTCCGAGAGCTCCAGAGAAATCTTCCGAGAGATCATGCCCAAGAGCCAGCAGAGAGAA 360
QY 361 AATGGTTTGAAGGAAGTGCAGAGGATCTGGCCCAAAATGATGGGAAACAGCTGTGC 420
DB 361 AATGGTTTGAAGGAAGTGCAGAGGATCTGGCCCAAAATGATGGGAAACAGCTGTGC 420
QY 421 CCCCCGGGAATCCAAAGTACCTTGGAGAGATTAACAAGACATCTGGACCCCAAAAGGGGG 480
DB 421 CCCCCGGGAATCCAAAGTACCTTGGAGAGATTAACAAGACATCTGGACCCCAAAAGGGGG 480
QY 481 AAACATGCTTGGACCCACAGACTGCGTGAGAGAAAGCAGCTGGTGGTTTATGAAGATC 540
DB 481 AAACATGCTTGGACCCACAGACTGCGTGAGAGAAAGCAGCTGGTGGTTTATGAAGATC 540
QY 541 AGCGACCTGAGGAAGATGACGAGTAACCTCCCTCG 576
DB 541 AGCGACCTGAGGAAGATGACGAGTAACCTCCCTCG 576
```

```
RESULT 2
US-09-833-039-5
; Sequence 5, Application US/09833039
; Publication No. US20030175960A1
; GENERAL INFORMATION:
; APPLICANT: Tureci, Ozlem
; APPLICANT: Sahin, Ugur
; APPLICANT: Pfreundschuh, Michael
; TITLE OF INVENTION: Tumor Associated Peptide and Uses Thereof
; FILE REFERENCE: LUD 5622.1
; CURRENT APPLICATION NUMBER: US/09/833,039
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: US 09/409,455
; PRIOR FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: US 09/344,040
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: US 09/105,839
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: US 08/851,130
; PRIOR FILING DATE: 1997-05-05
```

```
; NUMBER OF SEQ ID NOS: 129
; SEQ ID NO 5
; LENGTH: 576
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-833-039-5

Query Match      100.0%; Score 576; DB 10; Length 576;
Best Local Similarity 100.0%; Pred. No. 2.8e-174;
Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAACGGAGACGACGCTTTGCAAGAGACCCAGGGATGATGCTCAAAATATCAGAGAAG 60
DB 1 ATGAACGGAGACGACGCTTTGCAAGAGACCCAGGGATGATGCTCAAAATATCAGAGAAG 60
QY 61 TTACGAAAGGCCCTTCGATGATATGTCCTTAAGAAAGAGTGGGAAAGATG 120
DB 61 TTACGAAAGGCCCTTCGATGATATGTCCTTAAGAAAGAGTGGGAAAGATG 120
QY 121 AAATCCTCGAGAAATCGTCTATGTATATGAAGCTAACTATGAGTCAATGACTAAA 180
DB 121 AAATCCTCGAGAAATCGTCTATGTATATGAAGCTAACTATGAGTCAATGACTAAA 180
QY 181 CTAGGTTTCAAGGTCACCTCCACCTTTTCATGCTAGTAAACGGGCTGCAGACTTCCAC 240
DB 181 CTAGGTTTCAAGGTCACCTCCACCTTTTCATGCTAGTAAACGGGCTGCAGACTTCCAC 240
QY 241 GGGAAATGATTTGGTAAAGATCGAAGACCAAGGAATCAGGTTTCAAGCTTCTCAGATGACT 300
DB 241 GGGAAATGATTTGGTAAAGATCGAAGACCAAGGAATCAGGTTTCAAGCTTCTCAGATGACT 300
QY 301 TTCCGAGAGCTCCAGAGAAATCTTCCGAGAGATCATGCCCAAGAGCCAGCAGAGAGAA 360
DB 301 TTCCGAGAGCTCCAGAGAAATCTTCCGAGAGATCATGCCCAAGAGCCAGCAGAGAGAA 360
QY 361 AATGGTTTGAAGGAAGTGCAGAGGATCTGGCCCAAAATGATGGGAAACAGCTGTGC 420
DB 361 AATGGTTTGAAGGAAGTGCAGAGGATCTGGCCCAAAATGATGGGAAACAGCTGTGC 420
QY 421 CCCCCGGGAATCCAAAGTACCTTGGAGAGATTAACAAGACATCTGGACCCCAAAAGGGGG 480
DB 421 CCCCCGGGAATCCAAAGTACCTTGGAGAGATTAACAAGACATCTGGACCCCAAAAGGGGG 480
QY 481 AAACATGCTTGGACCCACAGACTGCGTGAGAGAAAGCAGCTGGTGGTTTATGAAGATC 540
DB 481 AAACATGCTTGGACCCACAGACTGCGTGAGAGAAAGCAGCTGGTGGTTTATGAAGATC 540
QY 541 AGCGACCTGAGGAAGATGACGAGTAACCTCCCTCG 576
DB 541 AGCGACCTGAGGAAGATGACGAGTAACCTCCCTCG 576
```

```
RESULT 3
US-10-177-277-5
; Sequence 5, Application US/10177277
; Publication No. US20030185844A1
; GENERAL INFORMATION:
; APPLICANT: Tureci, Ozlem
; APPLICANT: Sahin, Ugur
; APPLICANT: Pfreundschuh, Michael
; APPLICANT: Ramensee, Hans Georg
; APPLICANT: Stevanovic, Stefan
; TITLE OF INVENTION: Method for Determining Presence of Cancer in a Sample By Deter
; TITLE OF INVENTION: Expression of an SSX Gene, Peptides Derived From Said SSX Gen
; FILE REFERENCE: LUD 5556.1
; CURRENT APPLICATION NUMBER: US/10/177,277
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US/09/344,040
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: US 09/105,839
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: US 08/851,130
```

;; PRIOR FILING DATE: 1997-05-05
;; NUMBER OF SEQ ID NOS: 132
;; SEQ ID NO 5
;; LENGTH: 576
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-10-177-277-5

Query Match 100.0%; Score 576; DB 14; Length 576;
Best Local Similarity 100.0%; Pred. No. 2.8e-174;
Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAACGAGAGCAGCGCTTTCAGAGGAGACCCAGGATGATGCTCAAAATATCAGAGAG 60
DB 1 ATGAACGAGAGCAGCGCTTTCAGAGGAGACCCAGGATGATGCTCAAAATATCAGAGAG 60
QY 61 TTACGAAAGGCTTCGATGATATTCGCAAAATACTTCTTAAGAAGAGTGGGAAAGATG 120
DB 61 TTACGAAAGGCTTCGATGATATTCGCAAAATACTTCTTAAGAAGAGTGGGAAAGATG 120
QY 121 AAATCCTCGGAGAAATCGTCTATGTATGATGAAGCTAAACTATGAGGTCACTACTAAA 180
DB 121 AAATCCTCGGAGAAATCGTCTATGTATGATGAAGCTAAACTATGAGGTCACTACTAAA 180
QY 181 CTAGGTTTCAAGGTCACCTCCACCTTTTCATGCGTAGTAAACGGCTGCAGACTTCCAC 240
DB 181 CTAGGTTTCAAGGTCACCTCCACCTTTTCATGCGTAGTAAACGGCTGCAGACTTCCAC 240
QY 241 GGGAAATGATTTTGGTAAACGATCGAAACACAGGAATCAGGTTGAACGTCCTCAGATGACT 300
DB 241 GGGAAATGATTTTGGTAAACGATCGAAACACAGGAATCAGGTTGAACGTCCTCAGATGACT 300
QY 301 TTCGACGCTCCAGAGATCTTCCGAGATCATGCCAGAGATCATGCCAGAGCAGAGAGAGAA 360
DB 301 TTCGACGCTCCAGAGATCTTCCGAGATCATGCCAGAGATCATGCCAGAGCAGAGAGAGAA 360
QY 361 AATGTTTGAAGGAAGTGCAGAGGATCTGCGCCACACAAAATGATGGAAACAGCTGTGC 420
DB 361 AATGTTTGAAGGAAGTGCAGAGGATCTGCGCCACACAAAATGATGGAAACAGCTGTGC 420
QY 421 CCCCAGGAAATCCAAAGTACCTTGGAGAGATTAACAAGATTAACAAGATTAACAAGAGGGG 480
DB 421 CCCCAGGAAATCCAAAGTACCTTGGAGAGATTAACAAGATTAACAAGATTAACAAGAGGGG 480
QY 481 AAACATGCTGCAGCCACACAGATCTGCGTAGAGAAAGCAGCTGGTGTATTAAGAGATC 540
DB 481 AAACATGCTGCAGCCACACAGATCTGCGTAGAGAAAGCAGCTGGTGTATTAAGAGATC 540
QY 541 AGCGACCTTGAGGAAGATGACGATTAACCTCCCTCG 576
DB 541 AGCGACCTTGAGGAAGATGACGATTAACCTCCCTCG 576

RESULT 4
US-10-117-937-599
; Sequence 599, Application US/10117937
; Publication No. US20030220239A1
; GENERAL INFORMATION:
; APPLICANT: CTL IMMUNO THERAPIES CORP.
; APPLICANT: SIMARD, John, J.L.
; APPLICANT: DIAMOND, David, C.
; APPLICANT: LIU, Liping
; APPLICANT: XIE, Zhidong
; TITLE OF INVENTION: EPITOPE SEQUENCES
; FILE REFERENCE: CILMM.027A
; CURRENT APPLICATION NUMBER: US/10/117,937
; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: US 60/282,211
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: US 60/337,017
; PRIOR FILING DATE: 2001-11-07
; PRIOR APPLICATION NUMBER: US 60/363,210
; PRIOR FILING DATE: 2002-03-07

;; NUMBER OF SEQ ID NOS: 602
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 599
;; LENGTH: 576
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-10-117-937-599

Query Match 100.0%; Score 576; DB 15; Length 576;
Best Local Similarity 100.0%; Pred. No. 2.8e-174;
Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAACGAGAGCAGCGCTTTCAGAGGAGACCCAGGATGATGCTCAAAATATCAGAGAG 60
DB 1 ATGAACGAGAGCAGCGCTTTCAGAGGAGACCCAGGATGATGCTCAAAATATCAGAGAG 60
QY 61 TTACGAAAGGCTTCGATGATATTCGCAAAATACTTCTTAAGAAGAGTGGGAAAGATG 120
DB 61 TTACGAAAGGCTTCGATGATATTCGCAAAATACTTCTTAAGAAGAGTGGGAAAGATG 120
QY 121 AAATCCTCGGAGAAATCGTCTATGTATGATGAAGCTAAACTATGAGGTCACTACTAAA 180
DB 121 AAATCCTCGGAGAAATCGTCTATGTATGATGAAGCTAAACTATGAGGTCACTACTAAA 180
QY 181 CTAGGTTTCAAGGTCACCTCCACCTTTTCATGCGTAGTAAACGGCTGCAGACTTCCAC 240
DB 181 CTAGGTTTCAAGGTCACCTCCACCTTTTCATGCGTAGTAAACGGCTGCAGACTTCCAC 240
QY 241 GGGAAATGATTTTGGTAAACGATCGAAACACAGGAATCAGGTTGAACGTCCTCAGATGACT 300
DB 241 GGGAAATGATTTTGGTAAACGATCGAAACACAGGAATCAGGTTGAACGTCCTCAGATGACT 300
QY 301 TTCGACGCTCCAGAGATCTTCCGAGATCATGCCAGAGATCATGCCAGAGCAGAGAGAGAA 360
DB 301 TTCGACGCTCCAGAGATCTTCCGAGATCATGCCAGAGATCATGCCAGAGCAGAGAGAGAA 360
QY 361 AATGTTTGAAGGAAGTGCAGAGGATCTGCGCCACACAAAATGATGGAAACAGCTGTGC 420
DB 361 AATGTTTGAAGGAAGTGCAGAGGATCTGCGCCACACAAAATGATGGAAACAGCTGTGC 420
QY 421 CCCCAGGAAATCCAAAGTACCTTGGAGAGATTAACAAGATTAACAAGATTAACAAGAGGGG 480
DB 421 CCCCAGGAAATCCAAAGTACCTTGGAGAGATTAACAAGATTAACAAGATTAACAAGAGGGG 480
QY 481 AAACATGCTGCAGCCACACAGATCTGCGTAGAGAAAGCAGCTGGTGTATTAAGAGATC 540
DB 481 AAACATGCTGCAGCCACACAGATCTGCGTAGAGAAAGCAGCTGGTGTATTAAGAGATC 540
QY 541 AGCGACCTTGAGGAAGATGACGATTAACCTCCCTCG 576
DB 541 AGCGACCTTGAGGAAGATGACGATTAACCTCCCTCG 576

RESULT 5
US-09-849-602-13
; Sequence 13, Application US/09849602
; Publication No. US20030165834A1
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew J.
; APPLICANT: Old, Lloyd J.
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Chen, Yao-Tseng
; TITLE OF INVENTION: Colon Cancer Antigen Panel
; FILE REFERENCE: L0461/7105(JRV)
; CURRENT APPLICATION NUMBER: US/09/849,602
; CURRENT FILING DATE: 2001-05-04
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 13
; LENGTH: 766
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-849-602-13

Query Match 86.2%; Score 496.6; DB 10; Length 766;
 Best Local Similarity 91.5%; Pred. No. 1e-148;
 Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 1 ATGAACGGAGACGACGCTTTCAGAGGAGACCCAGGATGCTCAAAATATCAGAGAAG 60
 DB 92 ATGAACGGAGACGACGCTTTCAGAGGAGACCCAGGATGCTCAAAATATCAGAGAAG 151

QY 61 TTACGAAAGGCTTCGATGATATTCGCAAAATCTCTCTAAGAAAGAGTGGGAAAGATG 120
 DB 152 ATCAAAAGGCTTCGATGATATTCGCAAAATCTCTCTAAGAAAGAGTGGGAAAGATG 211

QY 121 AAATCTCGAGAAATCGTCTATGTGTATATGAAGCTAAACTATGAGGTCTAGCTAAA 180
 DB 212 AAAGCTTCGAGAAATCTCTATGTGTATATGAAGAGAAAGTATGAGGCTATGACTAAA 271

QY 181 CTAGGTTTCAAGTCAACCTTCCTCCACCTTCCTCCAGTCTAGTAAACGGCTGCAGACTTCCAC 240
 DB 272 CTAGGTTTCAAGGCTCCACCTTCCTCCACCTTCCTCCAGTCTAGTAAACGGCTGCAGACTTCCAG 331

QY 241 GGAATGATTTGGTAAACGATCGAATCAACCTTCCTCCAGTCTAGTAAACGGCTGCAGACTTCCAG 300
 DB 332 GGAATGATTTGGTAAACGATCGAATCAACCTTCCTCCAGTCTAGTAAACGGCTGCAGACTTCCAG 391

QY 301 TTGGCAGCTCCAGAGATCTTCCGAGAGATCATGCCCAAGAGCCAGCAGAGGAGAA 360
 DB 392 TTGGCAGCTCCAGAGATCTTCCGAGAGATCATGCCCAAGAGCCAGCAGAGGAGAA 451

QY 361 AATGTTTGAAGAAAGTCCAGAGGATCTGGCCCAACAAATCATGGAAACAGCTGTGC 420
 DB 452 AATGTTTGAAGAAAGTCCAGAGGATCTGGCCCAACAAATCATGGAAACAGCTGTGC 511

QY 421 CCCCCGGGAAATCCAACTACTCTGAGAGATTAACAAGACATCTGGACCCAAAGGGGG 480
 DB 512 CCCCCGGGAAATCCAACTACTCTGAGAGATTAACAAGACATCTGGACCCAAAGGGGG 571

QY 541 AGGACCTCGAGAGATGACGAGTAACCTCCCTC 575
 DB 632 AGGACCTCGAGAGATGACGAGTAACCTCCCTC 666

RESULT 6
 US-10-207-655-84
 ; Sequence 84, Application US/10207655
 ; Publication No. US20030118592A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ledbetter, Jeffrey A.
 ; APPLICANT: Hayden-Ledbetter, Martha S.
 ; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS
 ; FILE REFERENCE: 39069.401C1
 ; CURRENT APPLICATION NUMBER: US/10/207, 655
 ; PRIOR FILING DATE: 2002-07-25
 ; NUMBER OF SEQ ID NOS: 426
 ; SOFTWARE: Patent in version 3.0
 ; SEQ ID NO 84
 ; LENGTH: 766
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-207-655-84

Query Match 86.2%; Score 496.6; DB 14; Length 766;
 Best Local Similarity 91.5%; Pred. No. 1e-148;
 Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 1 ATGAACGGAGACGACGCTTTCAGAGGAGACCCAGGATGATGCTCAAAATATCAGAGAAG 60
 DB 92 ATGAACGGAGACGACGCTTTCAGAGGAGACCCAGGATGATGCTCAAAATATCAGAGAAG 151

QY 61 TTACGAAAGGCTTCGATGATATTCGCAAAATCTCTCTAAGAAAGAGTGGGAAAGATG 120
 DB 152 ATCAAAAGGCTTCGATGATATTCGCAAAATCTCTCTAAGAAAGAGTGGGAAAGATG 211

QY 121 AAATCTCGAGAAATCGTCTATGTGTATATGAAGCTAAACTATGAGGTCTAGCTAAA 180
 DB 212 AAAGCTTCGAGAAATCTCTATGTGTATATGAAGAGAAAGTATGAGGCTATGACTAAA 271

QY 181 CTAGGTTTCAAGTCAACCTTCCTCCACCTTCCTCCAGTCTAGTAAACGGCTGCAGACTTCCAC 240
 DB 272 CTAGGTTTCAAGGCTCCACCTTCCTCCACCTTCCTCCAGTCTAGTAAACGGCTGCAGACTTCCAG 331

QY 241 GGAATGATTTGGTAAACGATCGAATCAACCTTCCTCCAGTCTAGTAAACGGCTGCAGACTTCCAG 300
 DB 332 GGAATGATTTGGTAAACGATCGAATCAACCTTCCTCCAGTCTAGTAAACGGCTGCAGACTTCCAG 391

QY 301 TTGGCAGCTCCAGAGATCTTCCGAGAGATCATGCCCAAGAGCCAGCAGAGGAGAA 360
 DB 392 TTGGCAGCTCCAGAGATCTTCCGAGAGATCATGCCCAAGAGCCAGCAGAGGAGAA 451

QY 361 AATGTTTGAAGAAAGTCCAGAGGATCTGGCCCAACAAATCATGGAAACAGCTGTGC 420
 DB 452 AATGTTTGAAGAAAGTCCAGAGGATCTGGCCCAACAAATCATGGAAACAGCTGTGC 511

QY 421 CCCCCGGGAAATCCAACTACTCTGAGAGATTAACAAGACATCTGGACCCAAAGGGGG 480
 DB 512 CCCCCGGGAAATCCAACTACTCTGAGAGATTAACAAGACATCTGGACCCAAAGGGGG 571

QY 481 AAACATGCTCGGACCCACAGACTCGGTGAGAGAAAGCAGCTGGTGTATGAGAGATC 540
 DB 572 GAACATGCTCGGACCCACAGACTCGGTGAGAGAAAGCAGCTGGTGTATGAGAGATC 631

QY 541 AGGACCTCGAGAGATGACGAGTAACCTCCCTC 575
 DB 632 AGGACCTCGAGAGATGACGAGTAACCTCCCTC 666

RESULT 7
 US-10-117-937-6
 ; Sequence 6, Application US/10117937
 ; Publication No. US2003022039A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CTL IMMUNO THERAPIES CORP.
 ; APPLICANT: SIMARD, John, J.L.
 ; APPLICANT: DIAMOND, David, C.
 ; APPLICANT: LIU, Liping
 ; APPLICANT: XIE, Zhidong
 ; TITLE OF INVENTION: EPITOPE SEQUENCES
 ; FILE REFERENCE: CTIMM.027A
 ; CURRENT APPLICATION NUMBER: US/10/117,937
 ; PRIOR FILING DATE: 2002-04-04
 ; PRIOR APPLICATION NUMBER: US 60/282,211
 ; PRIOR FILING DATE: 2001-04-06
 ; PRIOR APPLICATION NUMBER: US 60/337,017
 ; PRIOR FILING DATE: 2001-11-07
 ; PRIOR APPLICATION NUMBER: US 60/363,210
 ; PRIOR FILING DATE: 2002-03-07
 ; NUMBER OF SEQ ID NOS: 602
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 6
 ; LENGTH: 766
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-117-937-6

Query Match 86.2%; Score 496.6; DB 15; Length 766;
 Best Local Similarity 91.5%; Pred. No. 1e-148;
 Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 1 ATGAACGGAGACGACGCTTTCAGAGGAGACCCAGGATGATGCTCAAAATATCAGAGAAG 60
 DB 92 ATGAACGGAGACGACGCTTTCAGAGGAGACCCAGGATGATGCTCAAAATATCAGAGAAG 151

QY 61 TTACGAAGGCGCTTCGATGATATTCGCAAAATCTCTCTTAAGAAAGAGTGGGAAAAGATG 120
 Db 152 ATCCAAAAGGCGCTTCGATGATATTCGCAAAATCTCTCTTAAGAAAGAGTGGGAAAAGATG 211
 QY 121 AAATCCTCGAGAAAATCGTCTATGTATATGAAGCTAAACTATGAGGTCTATGACTAAA 180
 Db 212 AAAGCCTCGAGAAAATCTTCTATGTATATGAAGAAAGATGATGAGGCTATGACTAAA 271
 QY 181 CTAGGTTTCAAGGTCAACCTCCACCTTCATGCTGTATGTAAGCTAAAGCTGCAAGTTCAC 240
 Db 272 CTAGGTTTCAAGGCCACCTCCACCTTCATGCTGTATGTAAGCTAAAGCTGCAAGTTCAC 331
 QY 241 GGAATGATTTTGGTAAAGCATGAAACACACAGGAATCAGGTTTGAACGCTCTCAGATGACT 300
 Db 332 GGAATGATTTGGTAATGACCTTAACCGTGGGAATCAGGTTTGAACGCTCTCAGATGACT 391
 QY 301 TTGGGAGCTCCAGAAATCTTCCGAAGATCATGCCCCAAGAACCCAGAGAGAGAA 360
 Db 392 TTGGGAGCTCCAGGAATCTCCCGAAGATCATGCCCCAAGAACCCAGAGAGAGAA 451
 QY 361 AATGTTTGAAGGAAGTGCAGAGCATCTGCCCAACAAATGATGGGAACAGCTGTGC 420
 Db 452 AATGATTTGGAGGAAGTGCAGAGCATCTGCCCAACAAATGATGGGAAGAGCTGTGC 511
 QY 421 CCCCAGGAAATCCAAAGTACCTTGGAGAGATTTAAACAGACATCTGGACCCCAAGAGGGG 480
 Db 512 CCCCAGGAAATCCAAAGTACCTTGGAGAGATTTAAACAGACATCTGGACCCCAAGAGGGG 571
 QY 481 AAACATGCTGACCCACAGACTGCGTGAGAGAAAGAGCTGCTGTTTATGAGAGATC 540
 Db 572 GAACATGCTGACCCACAGACTGCGTGAGAGAAAGAGCTGCTGTTTATGAGAGATC 631
 QY 541 AGCGACCTTGAGGAAGATGACGAGTAACCTCCCTC 575
 Db 632 AGCGACCTTGAGGAAGATGACGAGTAACCTCCCTC 666

RESULT 8

US-09-833-039-2
 ; Sequence 2, Application US/09833039
 ; Publication No. US20030175960A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Tureci, Ozlem
 ; APPLICANT: Sahin, Ugur
 ; APPLICANT: Pfreundschuh, Michael
 ; TITLE OF INVENTION: Tumor Associated Peptide and Uses Thereof
 ; FILE REFERENCE: LUD 5622.1
 ; CURRENT APPLICATION NUMBER: US/09/833,039
 ; CURRENT FILING DATE: 2001-04-12
 ; PRIOR APPLICATION NUMBER: US/09/409,455
 ; PRIOR FILING DATE: 1999-09-30
 ; PRIOR APPLICATION NUMBER: US/09/344,040
 ; PRIOR FILING DATE: 1999-06-25
 ; PRIOR APPLICATION NUMBER: US/09/105,839
 ; PRIOR FILING DATE: 1998-06-26
 ; PRIOR APPLICATION NUMBER: US/08/851,130
 ; PRIOR FILING DATE: 1997-05-05
 ; NUMBER OF SEQ ID NOS: 129
 ; SEQ ID NO 2
 ; LENGTH: 931
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-833-039-2

Query Match 86.2%; Score 496.6; DB 10; Length 931;
 Best Local Similarity 91.5%; Pred. No. 1.1e-148;
 Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
 QY 1 ATCAACGGAGACGACGCTTTGCAAGGAGACCCAGGATGATGCTCAAAATATCAGAGAAG 60
 Db 99 ATGAAAGGAGACGACGCTTTGCAAGGAGACCCAGGATGATGCTCAAAATATCAGAGAAG 158
 QY 61 TTACGAAGGCGCTTCGATGATATTCGCAAAATCTCTCTTAAGAAAGAGTGGGAAAAGATG 120

Db 159 ATCCAAAAGGCGCTTCGATGATATTCGCAAAATCTCTCTTAAGAAAGAGTGGGAAAAGATG 218
 QY 121 AAATCCTCGAGAAAATCGTCTATGTATATGAAGCTAAACTATGAGGTCTATGACTAAA 180
 Db 219 AAAGCCTCGAGAAAATCTTCTATGTATATGAAGAAAGATGATGAGGCTATGACTAAA 278
 QY 181 CTAGGTTTCAAGGTCAACCTCCACCTTCATGCTGTATGTAAGCTAAAGCTGCAAGTTCAC 240
 Db 279 CTAGGTTTCAAGGCCACCTCCACCTTCATGCTGTATGTAAGCTAAAGCTGCAAGTTCAC 338
 QY 241 GGAATGATTTTGGTAAAGCATGAAACACACAGGAATCAGGTTTGAACGCTCTCAGATGACT 300
 Db 339 GGAATGATTTGGTAATGACCTTAACCGTGGGAATCAGGTTTGAACGCTCTCAGATGACT 398
 QY 301 TTGGGAGCTCCAGAAATCTTCCGAAGATCATGCCCCAAGAACCCAGAGAGAGAA 360
 Db 399 TTGGGAGCTCCAGGAATCTCCCGAAGATCATGCCCCAAGAACCCAGAGAGAGAA 458
 QY 361 AATGTTTGAAGGAAGTGCAGAGCATCTGCCCAACAAATGATGGGAACAGCTGTGC 420
 Db 459 AATGATTTGGAGGAAGTGCAGAGCATCTGCCCAACAAATGATGGGAAGAGCTGTGC 518
 QY 421 CCCCAGGAAATCCAAAGTACCTTGGAGAGATTTAAACAGACATCTGGACCCCAAGAGGGG 480
 Db 519 CCCCAGGAAATCCAAAGTACCTTGGAGAGATTTAAACAGACATCTGGACCCCAAGAGGGG 578
 QY 481 AAACATGCTGACCCACAGACTGCGTGAGAGAAAGAGCTGCTGTTTATGAGAGATC 540
 Db 579 GAACATGCTGACCCACAGACTGCGTGAGAGAAAGAGCTGCTGTTTATGAGAGATC 638
 QY 541 AGCGACCTTGAGGAAGATGACGAGTAACCTCCCTC 575
 Db 639 AGCGACCTTGAGGAAGATGACGAGTAACCTCCCTC 673

RESULT 9

US-10-177-277-2
 ; Sequence 2, Application US/10177277
 ; Publication No. US2003018584A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Tureci, Ozlem
 ; APPLICANT: Sahin, Ugur
 ; APPLICANT: Pfreundschuh, Michael
 ; APPLICANT: Ramensee, Hans Georg
 ; APPLICANT: Stevanovic, Stefan
 ; TITLE OF INVENTION: Method for Determining Presence of Cancer in a Sample By Determining Presence of an SSX Gene, Peptides Derived From Said SSX Gene
 ; TITLE OF INVENTION: Expression of an SSX Gene, and Uses Thereof
 ; FILE REFERENCE: LUD 5556.1
 ; CURRENT APPLICATION NUMBER: US/10/177,277
 ; CURRENT FILING DATE: 2002-06-21
 ; PRIOR APPLICATION NUMBER: US/09/344,040
 ; PRIOR FILING DATE: 1999-06-25
 ; PRIOR APPLICATION NUMBER: US/09/105,839
 ; PRIOR FILING DATE: 1998-06-26
 ; PRIOR APPLICATION NUMBER: US/08/851,130
 ; PRIOR FILING DATE: 1997-05-05
 ; NUMBER OF SEQ ID NOS: 132
 ; SEQ ID NO 2
 ; LENGTH: 931
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-177-277-2

Query Match 86.2%; Score 496.6; DB 14; Length 931;
 Best Local Similarity 91.5%; Pred. No. 1.1e-148;
 Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
 QY 1 ATCAACGGAGACGACGCTTTGCAAGGAGACCCAGGATGATGCTCAAAATATCAGAGAAG 60
 Db 99 ATGAAAGGAGACGACGCTTTGCAAGGAGACCCAGGATGATGCTCAAAATATCAGAGAAG 158

QY 61 TTACGAAAGCCCTTCGATGATATGCAAAATACCTCTTAAGAAAGAGTGGAAAGATG 120
 Db 159 ATCCAAAGGCCCTTCGATGATATGCAAAATACCTCTTAAGAAAGAGTGGAAAGATG 218
 QY 121 AATCTCTCGAGAAATCGTCTATGTATATGAAGTAACTATGAGGTCTGACTAAA 180
 Db 219 AAGCCCTCGAGAAATCTCTATGTATATGAAGAAAGTATGAGGCTATGACTAAA 278
 QY 181 CTAGGTTCAAGTCAACCTCCACCTTCCAGGATCATGCCGCTCCAGTCCAC 240
 Db 279 CTAGGTTCAAGGCCACCTCCACCTTCCAGGATCATGCCGCTCCAGTCCAC 338
 QY 241 GGAATGATTTGGTAAGTCAACCTCCAGGATCATGCCGCTCCAGTCCAC 300
 Db 339 GGAATGATTTGGTAAGTCAACCTCCAGGATCATGCCGCTCCAGTCCAC 398
 QY 301 TTCCGAGCCCTCCAGAAATCTCTCCGAGATCATGCCGCTCCAGGATCATGCC 360
 Db 399 TTCCGAGCCCTCCAGAAATCTCTCCGAGATCATGCCGCTCCAGGATCATGCC 458
 QY 361 AATGGTTGAAGAGTCCAGAGGATCATGCCGCTCCAGGATCATGCCGCTCCAG 420
 Db 459 AATGATTCGAGAGTCCAGAGGATCATGCCGCTCCAGGATCATGCCGCTCCAG 518
 QY 421 CCCCAGGAAATCAAGTCAACCTCCAGGATCATGCCGCTCCAGGATCATGCC 480
 Db 519 CCCCAGGAAATCAAGTCAACCTCCAGGATCATGCCGCTCCAGGATCATGCC 578
 QY 481 AATGATTCGAGAGTCCAGAGGATCATGCCGCTCCAGGATCATGCCGCTCCAG 540
 Db 579 GAAATGCTCGAGAGTCCAGAGGATCATGCCGCTCCAGGATCATGCCGCTCCAG 638
 QY 541 AGCGACCTCGAGAGTCCAGAGGATCATGCCGCTCCAGGATCATGCCGCTCCAG 575
 Db 639 AGCGACCTCGAGAGTCCAGAGGATCATGCCGCTCCAGGATCATGCCGCTCCAG 673

RESULT 10

US-09-954-531-1000
 ; Sequence 1000, Application US/09954531
 ; Patent No. US20020165180A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Weaver, Zoe
 ; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand
 ; FILE REFERENCE: 689290-77
 ; CURRENT APPLICATION NUMBER: US/09/954,531
 ; CURRENT FILING DATE: 2002-05-02
 ; PRIOR FILING DATE: 2000-09-18
 ; PRIOR APPLICATION NUMBER: US/60/233,133
 ; PRIOR FILING DATE: 2000-09-20
 ; PRIOR APPLICATION NUMBER: US/60/234,009
 ; PRIOR FILING DATE: 2000-09-22
 ; PRIOR APPLICATION NUMBER: US/60/234,034
 ; PRIOR FILING DATE: 2000-09-22
 ; PRIOR APPLICATION NUMBER: US/60/234,509
 ; PRIOR FILING DATE: 2000-09-22
 ; PRIOR APPLICATION NUMBER: US/60/234,567
 ; NUMBER OF SEQ ID NOS: 1392
 ; SOFTWARE: Patent in version 3.0
 ; SEQ ID NO 1000
 ; LENGTH: 1309
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(1309)
 ; OTHER INFORMATION: n=a,t,g or c

Query Match 86.28; Score 496.6; DB 9; Length 1309;
 Best Local Similarity 91.5%; Pred. No. 1.4e-148;
 Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 1 ATGAACGAGACGACGCTTTGCAAGGAGACCAGGAGATGATCTCAATATCAGAGAAG 60
 Db 99 ATGAACGAGACGACGCTTTGCAAGGAGACCAGGAGATGATCTCAATATCAGAGAAG 158
 QY 61 TTACGAAGGCCCTTCGATGATATGCAAAATACCTCTTAAGAAAGAGTGGAAAGATG 120
 Db 159 ATCCAAAGGCCCTTCGATGATATGCAAAATACCTCTTAAGAAAGAGTGGAAAGATG 218
 QY 121 AATCTCTCGAGAAATCGTCTATGTATATGAAGTAACTATGAGGTCTGACTAAA 180
 Db 219 AAGCCCTCGAGAAATCTCTATGTATATGAAGAAAGTATGAGGCTATGACTAAA 278
 QY 181 CTAGGTTCAAGTCAACCTCCACCTTCCAGGATCATGCCGCTCCAGTCCAC 240
 Db 279 CTAGGTTCAAGGCCACCTCCACCTTCCAGGATCATGCCGCTCCAGTCCAC 338
 QY 241 GGAATGATTTGGTAAGTCAACCTCCAGGATCATGCCGCTCCAGTCCAC 300
 Db 339 GGAATGATTTGGTAAGTCAACCTCCAGGATCATGCCGCTCCAGTCCAC 398
 QY 301 TTCCGAGCCCTCCAGAAATCTCTCCGAGATCATGCCGCTCCAGGATCATGCC 360
 Db 399 TTCCGAGCCCTCCAGAAATCTCTCCGAGATCATGCCGCTCCAGGATCATGCC 458
 QY 361 AATGGTTGAAGAGTCCAGAGGATCATGCCGCTCCAGGATCATGCCGCTCCAG 420
 Db 459 AATGATTCGAGAGTCCAGAGGATCATGCCGCTCCAGGATCATGCCGCTCCAG 518
 QY 421 CCCCAGGAAATCAAGTCAACCTCCAGGATCATGCCGCTCCAGGATCATGCC 480
 Db 519 CCCCAGGAAATCAAGTCAACCTCCAGGATCATGCCGCTCCAGGATCATGCC 578
 QY 481 AATGATTCGAGAGTCCAGAGGATCATGCCGCTCCAGGATCATGCCGCTCCAG 540
 Db 579 GAAATGCTCGAGAGTCCAGAGGATCATGCCGCTCCAGGATCATGCCGCTCCAG 638
 QY 541 AGCGACCTCGAGAGTCCAGAGGATCATGCCGCTCCAGGATCATGCCGCTCCAG 575
 Db 639 AGCGACCTCGAGAGTCCAGAGGATCATGCCGCTCCAGGATCATGCCGCTCCAG 673

RESULT 11

US-09-954-531-1392
 ; Sequence 1392, Application US/09954531
 ; Patent No. US20020165180A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Weaver, Zoe
 ; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Ca
 ; FILE REFERENCE: 689290-77
 ; CURRENT APPLICATION NUMBER: US/09/954,531
 ; CURRENT FILING DATE: 2002-05-02
 ; PRIOR FILING DATE: 2000-09-18
 ; PRIOR APPLICATION NUMBER: US/60/233,133
 ; PRIOR FILING DATE: 2000-09-20
 ; PRIOR APPLICATION NUMBER: US/60/234,009
 ; PRIOR FILING DATE: 2000-09-22
 ; PRIOR APPLICATION NUMBER: US/60/234,034
 ; PRIOR FILING DATE: 2000-09-22
 ; PRIOR APPLICATION NUMBER: US/60/234,509
 ; PRIOR FILING DATE: 2000-09-22
 ; PRIOR APPLICATION NUMBER: US/60/234,567
 ; NUMBER OF SEQ ID NOS: 1392
 ; SOFTWARE: Patent in version 3.0
 ; SEQ ID NO 1392
 ; LENGTH: 1309
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(1309)
 ; OTHER INFORMATION: n=a,t,g or c

ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. US20030023057Alman D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5480
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 318-3168
TELEFAX: (212) 752-5958
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 576 nucleotides
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-975-856-2

Query Match 83.7%; Score 482.2; DB 10; Length 576;
Best Local Similarity 89.9%; Pred. No. 3.7e-144;
Matches 517; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 1 ATGAACGGAGACGCGCTTTCGAAGGAGACCCAGGGATGATGCTCAATATCAGAGAG 60
DB 1 ATGAACGGAGACGCGCTTTCGAAGGAGACCCAGGGATGATGCTCAATATCAGAGAG 60
QY 61 TTACGAAAGCGCTTCGATGATATGCCAAATATCTTCTTAAGAAAGAGTGGGAAAGATG 120
DB 61 ATGCAAAAGCGCTTCGATGATATGCCAAATATCTTCTTAAGAAAGAGTGGGAAAGATG 120
QY 121 AAATCTCGGAGAAATCGTCTATGTATATGAAGCTAACTATGAGGTGATGACTTAA 180
DB 121 AAAGCCTCGGAGAAATCATCTATGTATATGAAGAAAGATGATGAGGCCATGACTTAA 180
QY 181 CTAGGTTTCAAGTCAACCTCCACCTTTCATCGCTAGTAAACGGGCTGCAGACTTCCAC 240
DB 181 CTAGGTTTCAAGGACACCTCCACCTTTCATCGCTAGTAAACGGGCTGCAGACTTCCAG 240
QY 241 GGAATGATTTTGGTAAAGATCGAAACACAGGAATCAGGTTGAAGCTCTCAGATGACT 300
DB 241 GGAATGATTTTGGTAAAGATCGAAACACAGGAATCAGGTTGAAGCTCTCAGATGACT 300
QY 301 TTCCGAGGCTCCAGAGATCTTCCGAGAGATCATGCCAAGAGCCAGAGAGAGAA 360
DB 301 TTCCGAGGCTCCAGAGATCTTCCGAGAGATCATGCCGAGAGCCAGAGAGAGGA 360
QY 361 AATGTTTGAAGAAAGTGCAGAGGATCTGCCCAACAAATGATGGGAAACAGCTGTGC 420
DB 361 AATGATTCAAAGGAGTGCAGAGGATCTGCCCAACAAATGATGGGAAACAGCTGTGC 420
QY 421 CCCCAGGAAATCCCAAGTACCTTGGAGAGATTAAACAGACATCTGGACCCAAAGGGGG 480
DB 421 CCTCAGGAAACTAAATACCTCTGAGAGGTTTAAACAGACATCTGGACCCAAAGGGGG 480
QY 481 AAACATGCTGACCCACAGATCTGCTGAGAGAAAGCAGCTGGTGTATGAAGAGATC 540
DB 481 AAACATGCTGACCCACAGATCTGCTGAGAGAAAGCAGCTGGTGTATGAAGAGATC 540
QY 541 AGCGACCTGAGAGATCAGGATTAACCTCCCTC 575
DB 541 AGCGACCTGCGAAGATGACGATTAACCTCCCTC 575

RESULT 14
US-09-833-039-6
Sequence 6, Application US/09833039
Publication No. US20030175960A1
GENERAL INFORMATION:
APPLICANT: Tureci, Ozlem
APPLICANT: Sahin, Ugur
APPLICANT: Pfreundschuh, Michael
TITLE OF INVENTION: Tumor Associated Peptide and Uses Thereof
FILE REFERENCE: LUD 5622.1
CURRENT APPLICATION NUMBER: US/09/833,039
CURRENT FILING DATE: 2001-04-12

PRIOR APPLICATION NUMBER: US 09/409,455
PRIOR FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: US 09/344,040
PRIOR FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: US 09/105,839
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: US 08/851,130
PRIOR FILING DATE: 1997-05-05
NUMBER OF SEQ ID NOS: 129
SEQ ID NO 6
LENGTH: 576
TYPE: DNA
ORGANISM: Homo sapiens
US-09-833-039-6

Query Match 83.7%; Score 482.2; DB 10; Length 576;
Best Local Similarity 89.9%; Pred. No. 3.7e-144;
Matches 517; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 1 ATGAACGGAGACGCGCTTTCGAAGGAGACCCAGGGATGATGCTCAATATCAGAGAG 60
DB 1 ATGAACGGAGACGCGCTTTCGAAGGAGACCCAGGGATGATGCTCAATATCAGAGAG 60
QY 61 TTACGAAAGCGCTTCGATGATATGCCAAATATCTTCTTAAGAAAGAGTGGGAAAGATG 120
DB 61 ATGCAAAAGCGCTTCGATGATATGCCAAATATCTTCTTAAGAAAGAGTGGGAAAGATG 120
QY 121 AAATCTCGGAGAAATCGTCTATGTATATGAAGCTAACTATGAGGTGATGACTTAA 180
DB 121 AAAGCCTCGGAGAAATCATCTATGTATATGAAGAAAGATGATGAGGCCATGACTTAA 180
QY 181 CTAGGTTTCAAGTCAACCTCCACCTTTCATCGCTAGTAAACGGGCTGCAGACTTCCAC 240
DB 181 CTAGGTTTCAAGGACACCTCCACCTTTCATCGCTAGTAAACGGGCTGCAGACTTCCAG 240
QY 241 GGAATGATTTTGGTAAAGATCGAAACACAGGAATCAGGTTGAAGCTCTCAGATGACT 300
DB 241 GGAATGATTTTGGTAAAGATCGAAACACAGGAATCAGGTTGAAGCTCTCAGATGACT 300
QY 301 TTCCGAGGCTCCAGAGATCTTCCGAGAGATCATGCCAAGAGCCAGAGAGAGAA 360
DB 301 TTCCGAGGCTCCAGAGATCTTCCGAGAGATCATGCCGAGAGCCAGAGAGAGGA 360
QY 361 AATGTTTGAAGAAAGTGCAGAGGATCTGCCCAACAAATGATGGGAAACAGCTGTGC 420
DB 361 AATGATTCAAAGGAGTGCAGAGGATCTGCCCAACAAATGATGGGAAACAGCTGTGC 420
QY 421 CCCCAGGAAATCCCAAGTACCTTGGAGAGATTAAACAGACATCTGGACCCAAAGGGGG 480
DB 421 CCTCAGGAAACTAAATACCTCTGAGAGGTTTAAACAGACATCTGGACCCAAAGGGGG 480
QY 481 AAACATGCTGACCCACAGATCTGCTGAGAGAAAGCAGCTGGTGTATGAAGAGATC 540
DB 481 AAACATGCTGACCCACAGATCTGCTGAGAGAAAGCAGCTGGTGTATGAAGAGATC 540
QY 541 AGCGACCTGAGAGATCAGGATTAACCTCCCTC 575
DB 541 AGCGACCTGCGAAGATGACGATTAACCTCCCTC 575

RESULT 15
US-10-177-277-6
Sequence 6, Application US/10177277
Publication No. US20030185844A1
GENERAL INFORMATION:
APPLICANT: Tureci, Ozlem
APPLICANT: Sahin, Ugur
APPLICANT: Pfreundschuh, Michael
APPLICANT: Ramensee, Hans Georg
APPLICANT: Stevanovic, Stefan
TITLE OF INVENTION: Method for Determining Presence of Cancer in a Sample By Detect
TITLE OF INVENTION: Expression of an SSX Gene, Peptides Derived From Said SSX Gene
TITLE OF INVENTION: Gene, and Uses Thereof


```
FILE REFERENCE: IUD 5556.1
; CURRENT APPLICATION NUMBER: US/10/177,277
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US/09/344,040
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: US 09/105,839
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: US 08/851,130
; PRIOR FILING DATE: 1997-05-05
; NUMBER OF SEQ ID NOS: 132
; SEQ ID NO 6
; LENGTH: 576
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-177-277-6

Query Match      83.7%; Score 482.2; DB 14; Length 576;
Best Local Similarity 89.9%; Pred. No. 3.7e-144;
Matches 517; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 1 ATGAACGGAGACGACGCCCTTTGCAAGGAGACCCAGGGATGATGCTCAATATCAGAGAAG 60
Db 1 ATGAACGGAGACGACGCCCTTTGACGAGACCTAGGGTTGGTTCTCAATATCAGAGAAG 60
QY 61 TTACGAAAGCCCTTCGATGATATGCGCAAAATCTTCTTAAGAAAGATGGGAAAAGATG 120
Db 61 ATGCAAAAGGCCCTTCGATGATATGCGCAAAATCTTCTCTGAGAAAGATGGGAAAAGATG 120
QY 121 AAATCCTCGAGAAATCGTCTATGTATATGAAGCTAAACTATGAGTCAATGACTAAA 180
Db 121 AAAGCCTCGAGAAATCTATGTATATGAAGAAAGATATGAGGCCATGACTAAA 180
QY 181 CTAGGTTTCAAGTCAACCTCCACCTTTTCATGCGTAGTAAACGGGCTGCAGATCCAC 240
Db 181 CTAGGTTTCAAGGCCACCTCCACCTTTTCATGCGTAAATAAACGGGTGCAGATCCAG 240
QY 241 GGAATGATTTTGGTAAACGTCGAAACACAGGAATCAGGTTGAAAGTCTCAGATGACT 300
Db 241 GGAATGATTTTGGTAAACCTTAACCTGGAATCAGGTTGAAACATCCTCAGATGACT 300
QY 301 TTCGGCAGGCTCCAGAGAATCTTCCGAAAGATCATGCCCAAGAGCCAGCAGAGGAAGAA 360
Db 301 TTCGGCAGGCTCCAGGGAATCTTCCGAAAGATCAGCCCGAGAGCCAGCAGAGGAAGAA 360
QY 361 AATGTTTGAAGAAAGTGCAGAGGATCTGCGCCACAAATGATGGGAAACAGCTGTGC 420
Db 361 AATGATTCAAAGGGAGTGCAGAGGATCTGCGCCACAGCAATGGGAAACAGCTGTGC 420
QY 421 CCCCCGGGAAATCCAAAGTACCTTTGGAGAAAGATTAAACAGACATCTGGACCCAAAAGGGG 480
Db 421 CCTCAGGAAACTAATACTCTGAGAGGTTTACAGACATCTGGACCCAAAAGGGG 480
QY 481 AAACATGCTTGGACCCACAGACTGCGTGAGAGAAAGCAGCTGGTGGTTTATGAAGATC 540
Db 481 AAACATGCTTGGACCCACAGAGTGCCTGAGAGAAAGCACTGGTGGTTTATGAAGATC 540
QY 541 AGCGACCTTGAGGAAGATGACGAGTAATCTCCCTC 575
Db 541 AGCGACCTTGGGAAGATGACGAGTAATCTCCCTC 575
```

Search completed: March 31, 2004, 18:32:55
Job time : 402 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: March 31, 2004, 13:39:56 ; Search time 14 Seconds
(without alignments)
4284.630 Million cell updates/sec

Title: US-09-975-856-1
Perfect score: 1025
Sequence: 1 ATGACGAGACGACGCGCTT.....ATGACGAGTAACCTCCCTCG 576

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Zgapop 6.0 , Zgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 283362

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-Q=/cgn2.1/USPO.spool.p/US09975856/runat.31032004.132903.14738/app.query.fasta.1.775
-DB=SwissProt_42 -QFMT=fastan -SUFFIX=rsb -MINMATCH=0.1 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=BLOSUM62 -TRANS=human40.cdi -LISI=45
-DOCALIGN=200 -THR SCORE=500 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09975856.qcgn.1.16 @runat.31032004.132903.14738 -NCPU=3
-NO.MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1004	98.0	188	1 SX4 HUMAN	O60224 homo sapien
2	801	78.1	188	1 SX3 HUMAN	Q39909 homo sapien
3	798	77.9	188	1 SX2 HUMAN	Q16385 homo sapien
4	784	76.5	188	1 SX1 HUMAN	Q16384 homo sapien
5	780	76.1	188	1 SX5 HUMAN	O60225 homo sapien
6	113.5	11.1	289	1 ZN75 HUMAN	P51815 homo sapien
7	101	9.9	642	1 ZN75 HUMAN	Q8td17 homo sapien
8	94	9.2	595	1 Z317 HUMAN	Q96pg6 homo sapien
C 9	89.5	8.9	1505	1 PK3G RAT	O70173 rattus norv
C 10	88.5	8.8	531	1 FMO1 CANFA	Q95la2 canis fami
11	88	8.6	743	1 REP1 MOUSE	O54916 mus musculu
C 12	87.5	8.7	531	1 FMO1 PIG	P16549 sus scrofa
13	87	8.5	744	1 REP1 HUMAN	Q96d71 homo sapien
14	86.5	8.4	396	1 SX11 CHICK	P48435 Gallus gall
15	86.5	8.4	446	1 ZN38 HUMAN	P17036 homo sapien
16	86.5	8.4	488	1 ZP92 MOUSE	Q62396 mus musculu
17	86.5	8.4	572	1 TCI1 MOUSE	Q61751 mus musculu
18	86	8.4	480	1 WR61 ARATH	Q8vrv6 arabidopsis

ALIGNMENTS

RESULT 1

SSX4_HUMAN
ID SSX4_HUMAN STANDARD; PRT; 188 AA.

AC O60224; O9UTJ9;

DT 15-JUL-1999 (Rel. 38, Created)

DT 15-JUL-1999 (Rel. 38, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE SSX4 protein.

GN SSX4.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_Taxid=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=98021352; PubMed=9378559;

RA Gure A.O., Tuereci O., Sahin U., Tsang S., Scanlan M.J., Jager E.,

RA Knuth A., Pfreundschuh M., Old L.J., Chen Y.-T.;

RT "SSX: a multigene family with several members transcribed in normal

testis and human cancer";

RL Int. J. Cancer 72:963-971(1997).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=22388257; PubMed=12477932;

RA Krausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Donald M.P., Casavant T.L., Prange C.,

RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Bracken

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalon D.K., Muny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,

RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalios D.E.,

RA Schnerch A., Schein J.E., Jones S.O.M., Marra M.A.;

19	86	8.4	834	1 CASL_HUMAN	Q14511 homo sapien
20	85.5	8.3	680	1 Z334_HUMAN	Q5hcz1 homo sapien
21	85.5	8.3	738	1 ZN84_HUMAN	P15223 homo sapien
22	84.5	8.2	429	1 HISX_METJA	Q58851 methanococc
23	84.5	8.2	463	1 Z331_HUMAN	Q9rxk6 homo sapien
24	84.5	8.4	534	1 FMO1_RABIT	P17636 oryctolagus
25	84	8.2	524	1 Z479_HUMAN	Q96jca homo sapien
26	83.5	8.1	636	1 ZF90_MOUSE	Q61967 mus musculu
27	83.5	8.1	1102	1 ANP_HUMAN	Q9H2p0 homo sapien
28	83.5	8.1	1630	1 MSP1_PLAFLW	P04932 plasmodium
29	83.5	8.1	1639	1 MSP1_PLAFLW	P04933 plasmodium
30	82.5	8.0	553	1 Z324_HUMAN	O75467 homo sapien
31	82.5	8.0	643	1 ZN74_HUMAN	P16587 beetle wester
32	82.5	8.0	1035	1 RRPO_BWYVF	P09507 beetle wester
33	81.5	8.0	504	1 Z205_HUMAN	O95201 homo sapien
34	81.5	8.0	639	1 GLCX_SOYEN	P11827 glycine max
35	81.5	8.0	1395	1 SF41_YEAST	P38904 saccharomyc
36	81.5	8.0	1443	1 DPO3_MYCPN	P75080 mycoplasma
37	81	7.9	458	1 ZN19_HUMAN	P17023 homo sapien
38	81	7.9	751	1 Z337_HUMAN	Q9y3m9 homo sapien
C 39	81	8.0	933	1 PERT_CANFA	Q8hyb7 canis fami
40	80.5	7.9	439	1 ZN10_HUMAN	P21506 homo sapien
41	80.5	7.9	470	1 Z436_HUMAN	Q9c0f3 homo sapien
42	80.5	7.9	659	1 Z304_HUMAN	Q9hcx3 homo sapien
43	80.5	7.9	998	1 GTFI_HUMAN	P78347 h general t
C 44	79.5	7.9	531	1 FMO1_HUMAN	Q01740 homo sapien
45	79.5	7.8	626	1 Z471_HUMAN	Q9bxb2 homo sapien

RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 Proc. Natl. Acad. Sci. U.S.A. 99:16999-16903 (2002).
 [3]
 RN SEQUENCE OF 1-155 FROM N.A.
 RA Blechschmidt K., Nyakatura G., Strom T.M., Drescher B., Menzel U.,
 RA Meindl A., Rosenthal A.;
 .RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Could act as a modulator of transcription.
 CC -!- SIMILARITY: Belongs to the SSX family.
 CC -!- SIMILARITY: Contains 1 KRAB-related domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U0841; AAC05820.1; .
 DR EMBL; BC005325; AAH05325.1; .
 DR EMBL; AF196972; AAF06796.1; .
 DR Genew; HGNC:11338; SSX4.
 DR MIM; 300326; .
 DR InterPro; IPR001909; KRAB.
 DR InterPro; IPR003655; KRAB_related.
 DR Pfam; PF01352; KRAB; 1.
 DR SMART; SM00349; KRAB; 1.
 DR PROSITE; PS50806; KRAB_RELATED; 1.
 KW Multigene family; Transcription regulation.
 FT DOMAIN 20 83
 SQ SEQUENCE 188 AA; 21858 MW; 4ACA2A8737507AE5 CRC64;

Alignment Scores:
 Pred. No.: 1.37e-91 Length: 188
 Score: 1004.00 Matches: 188
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 97.95% Indels: 0
 DB: 1 Gaps: 0
 US-09-975-856-1 (1-576) x SSX4_HUMAN (1-188)
 QY 1 ATGACGGAGAGCGCTTTTCAGAGGAGACCCAGGATGCTCAATATCAGAGAG 60
 Db 1 MetAnGlyAspAlaPheAlaArgProArgAspAlaGlnIleSerGluLys 20
 QY 61 TTACGAAGGCTTCGATGATATCCCAATACCTTCTTAAGAAAGAGTGGGAAAGATG 120
 Db 21 LeuArgLysAlaPheAspAlaLysTyrPheSerLysLysGluTyrGluLysMet 40
 QY 121 AAATCTCGGAGAAATCGTCTATGCTATGATGAAGCTAACTAGGTCATGACTAAA 180
 Db 41 LysSerSerGluLysIleValTyrValTyrMetLysLeuAenTyrGluValMetThrLys 60
 QY 181 CTAGTTTCAAGGTCAACCTCCACCTTTCATGCTAGTAAACGGGCTGCAGACTTCCAC 240
 Db 61 LeuGlyPheLysValThrLeuProPheMetArgSerLysArgAlaAspPheHis 80
 QY 241 GGGATGATTTGGTAACGATCGAAACACAGGATCAGTTTGAAGCTTCAGTACT 300
 Db 81 GlyAenAppPheGlyAsnAspArgAsnHisArgAenGlnValGluArgProGlnMetThr 100
 QY 301 TTCGGCAGCTCCAGAGATCTTCCGAGATCATGCCAAGATCATGCCAAGACCCAGCAGAGAGAA 360
 Db 101 PheGlySerLeuGlnArgIlePheProLysIleMetProLysLysProAlaGluGlu 120
 QY 361 AATGTTTGAAGAGATGCTCCAGGAGCTTGGCCCAAAATGATGGAAACAGCTGTC 420
 Db 121 AsnGlyLeuLysGluValProGluAlaSerGlyProGlnAsnAspGlyLysGlnLeuLys 140
 QY 421 CCCCGGGAATCCCAAGTACCTTGGAGAGATTAACAAGACATCTGGACCCCAAGGGGG 480

Db 141 ProProGlyAsnProSerThrLeuGluLysIleAenLysThrSerGlyProLysArgGly 160
 QY 481 AAACATGCTGAGACCCACAGACTGCTGAGAGAAAGAGCTGCTGTTTATGAAGATC 540
 Db 161 LysHisAlaTyrThrHisArgLeuArgGluAenLysGlnLeuValValTyrGluGlu 180
 QY 541 ACGGACCTGAGAGAGATGACGAG 564
 Db 181 SerAspProGluGluAspAspGlu 188
 RESULT 2
 SSX3_HUMAN STANDARD; PRT; 188 AA.
 ID SSX3_HUMAN STANDARD; PRT; 188 AA.
 AC OS9979; O60223;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE SSX3 protein.
 GN SSX3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Fibroblasts, and Testis;
 RX MEDLINE=96302330; PubMed=8697803;
 RA de Leeuw B., Balemans M., Geurts van Kessel A.;
 RT "A novel Kruppel-associated box containing the SSX gene (SSX3) on the
 RT human X chromosome is not implicated in t(X;18)-positive synovial
 RT sarcomas.";
 RL Cytogenet. Cell Genet. 73:179-183 (1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98021352; PubMed=9378559;
 RA Gure A.O., Tuereci O., Sahin U., Tsang S., Scanlan M.J., Jager E.,
 RA Kruth A., Pfeundschnig M., Old L.J., Chen Y.-T.;
 RT "SSX: a multigene family with several members transcribed in normal
 RT testis and human cancer.";
 RL Int. J. Cancer 72:965-971 (1997).
 CC -!- FUNCTION: Could act as a modulator of transcription.
 CC -!- SIMILARITY: Belongs to the SSX family.
 CC -!- SIMILARITY: Contains 1 KRAB-related domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; S82471; AAB37436.2; .
 DR EMBL; U90840; AAC05819.1; .
 DR Genew; HGNC:11337; SSX3.
 DR MIM; 300325; .
 DR InterPro; IPR001909; KRAB.
 DR InterPro; IPR003655; KRAB_related.
 DR Pfam; PF01352; KRAB; 1.
 DR SMART; SM00349; KRAB; 1.
 DR PROSITE; PS50806; KRAB_RELATED; 1.
 KW Multigene family; Transcription regulation.
 FT DOMAIN 20 83
 FT CONFLICT 95 95 L -> Q (IN REF. 2).
 SQ SEQUENCE 188 AA; 21682 MW; 368C7487C1450E72 CRC64;

Alignment Scores:
 Pred. No.: 1.85e-71 Length: 188
 Score: 801.00 Matches: 152
 Percent Similarity: 84.57% Conservative: 7
 Best Local Similarity: 80.85% Mismatches: 29
 Query Match: 78.15% Indels: 0

```
DB: 1 1 Gaps: 0
US-09-975-856-1 (1-576) x SSX2_HUMAN (1-188)
QY 1- ATGAACGGAGACACCCCTTCGACGAGACACCCAGGATGATGCTCAATATCAGAGAG 60
Db 1- MCAASGLYASpAspThrPheAlaAlaGAPGProThrValGlyAlaGlnIleProGluLys 20
QY 61- TTACGAAGAGCCCTTCGATGATATTCGCAAAATCTCTCTAAGAAAGAGTGGGAAAGATG 120
Db 21- IIEGlnLYAlaPheAspAspIleAlaLYsTyPheSerLYsGlnGluTrpGluLysMet 40
QY 121- AATCTCCGAGAGAAATCTCTATGCTATATGAAGCTAACTAAGTATGAGTCTAGCTAA 180
Db 41- LysValSerGluLYsIleValTYrMetLYsArgLYsTYrGluAlaMetThrLYs 60
QY 181- CTAGGTTTCAAGGTCACCTCCACCTTCATCGCTAGTAAACGGGCTGCACACTTCAC 240
Db 61- LeuGlyPheLYsAlaIleLeuProSerPheMetArgAsnLYsArgValThrAspPheGln 80
QY 241- GGAATGATTTGGTAACGATCGAACCACAGGAATCAGGTTGAACGTCCTCAGATGACT 300
Db 81- GlyAsnAspPheAspAsnAspProAsnArgGlyAsnGlnValLeuArgProGlnMetThr 100
QY 301- TTCGGCAGCCTCCAGAGAAATCTTCCGGAAGATCATGCCCAAGAGCCAGCAGAGAGAA 360
Db 101- PheGlyArgLeuGlnGlyIlePheProLYsIleMetProLYsIleProIleGluGly 120
QY 361- ATGTTTGAAGAAATGTCAGAGGCAATTCGCCCAAAATATGATGGGAAACAGCTGTC 420
Db 121- AsnValSerLYsGluValProGluAlaSerGlyProGlnAsnAspGlyLYsGlnLeuCYs 140
QY 421- CCCCAGGAAATCCCAAGTACCTTGGAGAGATTAACAAGACATCTGCACCCAAAGGGGG 480
Db 141- ProProGlyLYsProThrThrSerGluLYsIleAsnMetIleSerGlyProLYsArgGly 160
QY 481- AAACATGCTGGACCCACAGATGCTGTCGAGAGAAACAGCTGCTGTATGAAGATC 540
Db 161- GluHisAlaIleTrpThrHisArgLeuArgGluArgGluLYsGlnLeuValIleTYrGluGluIle 180
QY 541- ACGACCCCTCAGAGAGATGACGAG 564
Db 181- SerAspProGluGluAspGlu 188

RESULT 3
SSX2_HUMAN STANDARD; PRT; 188 AA.
ID Q1635; Q16404; Q96IP7;
AC 15-DEC-1998 (Rel. 37, Created)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE SSX2 protein (synovial sarcoma, X breakpoint 2) (SSX) (HOM-MEL-40).
GN SSX2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=95292974; PubMed=7539744;
RA Crew A.J., Clark J., Fisher C., Gill S., Grimer R., Chand A.,
RA Shipley J., Gusterson B.A., Cooper C.S.;
RT "Fusion of SYT to two genes, SSX1 and SSX2, encoding proteins with
RT homology to the Kruppel-associated box in human synovial sarcoma.";
RL EMBL J. 14:2333-2340(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RX MEDLINE=22368257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uadin T.B., Tohiyuki S., Carninci P., Prange C.,
RA Raha S.S., Iqbalan N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Guadagnoli F.H.,
RA Richards S., Wozney K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Morley D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A.C., Shevchenko Y., Bouffard G.G.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.F., Skalska U., Smallus D.B.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP SEQUENCE OF 111-188 FROM N.A.
RC TISSUE=Synovial sarcoma;
RX MEDLINE=95384157; PubMed=7655467;
RA de Leeuw B., Bailemans M., Olde Weghuis D., Geurts van Kessel A.;
RT "Identification of two alternative fusion genes, SYT-SSX1 and
RT SYT-SSX2, in t(X;18)(p11.2;q11.2)-positive synovial sarcomas.";
RL Hum. Mol. Genet. 4:1097-1099(1995).
RN [4]
RP SEQUENCE OF 111-188 FROM N.A. (SSXT-SSX2 FUSION PROTEIN).
RC TISSUE=Synovial sarcoma;
RX MEDLINE=96094743; PubMed=7951320;
RA Clark J., Rocques P.J., Crew A.J., Gill S., Shipley J., Chan A.M.-L.,
RA Gusterson B.A., Cooper C.S.;
RT "Identification of novel genes, SYT and SSX, involved in the
RT t(X;18)(p11.2;q11.2) translocation found in human synovial sarcoma.";
RL Nat. Genet. 7:502-508(1994).
RN [5]
RP SEQUENCE OF 68-116 FROM N.A. (SSXT-SSX2 FUSION PROTEIN).
RX MEDLINE=96094743; PubMed=795284;
RA Fligman I., Lonardo F., Jhanwar S.C., Gerald W.L., Woodruff J.,
RA Ladanyi M.;
RT "Molecular diagnosis of synovial sarcoma and characterization of a
RT variant SYT-SSX2 fusion transcript.";
RL Am. J. Pathol. 147:1592-1599(1995).
CC -!- FUNCTION: Could act as a modulator of transcription.
CC -!- TISSUE SPECIFICITY: Expressed at high level in the testis.
CC Expressed at low level in thyroid. Not detected in tonsil, colon,
CC lung, spleen, prostate, kidney, striated and smooth muscles.
CC Detected in rhabdomyosarcoma and fibrosarcoma cell lines. Not
CC detected in mesenchymal and epithelial cell lines.
CC -!- DISEASE: Participates in a t(X;18)(p11.2;q11.2) chromosomal
CC translocation that is specifically found in more than 80% of
CC synovial sarcoma and produces the SSXT-SSX1 or SSXT-SSX2 fusion
CC products. These hybrid proteins are probably responsible for
CC transforming activity. Heterogeneity in the position of the
CC breakpoint can occur (low frequency).
CC -!- SIMILARITY: Belongs to the SSX family.
CC -!- SIMILARITY: Contains 1 KRAB-related domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X86175; CAA60111.1; --
DR EMBL; BC007343; AA07343.1; --
DR EMBL; BC016957; AA016957.1; --
DR EMBL; S79332; AAB35379.1; --
DR EMBL; X79200; --; NOT_ANNOTATED_CDS.
DR EMBL; S79894; AAB35674.1; ALT_INIT.
DR PIR; S55058; S55058.
DR Genew; HGNC:11336; SSX2.
```

DR MTM; 300192; ..
 DR GO; 0006355; P:regulation of transcription, DNA-dependent; NAS.
 DR InterPro; IPR001909; KRAB.
 DR InterPro; IPR003655; KRAB_related.
 DR Pfam; PF01352; KRAB; 1.
 DR SMART; SMO0349; KRAB; 1.
 DR PROSITE; PSS0806; KRAB RELATED; 1.
 DR Chromosomal translocation; Proto-oncogene; Multigene family;
 KW Transcription regulation.
 FT DOMAIN 20 83 KRAB-RELATED.
 FT SITE 68 69 BREAKPOINT FOR TRANSLOCATION TO FORM THE
 FT SITE 110 111 SSX1-SSX2 FUSION PROTEIN (RARE).
 FT SITE 169 169 BREAKPOINT FOR TRANSLOCATION TO FORM THE
 FT CONFLICT 169 169 SSX1-SSX2 FUSION PROTEIN.
 FT CONFLICT 169 169 R -> P (IN REF. 2; AAH07343).
 SQ SEQUENCE 188 AA; 21620 MW; BF5D18AA5F45B1B1 CRC64;

Alignment Scores:
 Pred. No.: 3.66e-71 Length: 188
 Score: 798.00 Matches: 150
 Percent Similarity: 86.17% Conservative: 12
 Best Local Similarity: 79.79% Mismatches: 26
 Query Match: 77.85% Indels: 0
 DB: 1 Gaps: 0

US-09-975-856-1 (1-576) x SSX2_HUMAN (1-188)

Qy 1 ATGAACGAGACGACGCTTTGCAAGGAGACCCAGGAGATGCTCAATATATCAGAGAG 60
 Db 1 MetAsnGlyAspAlaPheAlaArgProThrValGlyAlaGlnIleProGluLys 20
 Qy 61 TTACGAAGCCCTCGATATATGCAATATCTCTCAAGAAAGTGGGAAAGATG 120
 Db 21 IleGlnLysAlaPheAspAlaPheAlaLysPheSerLysGluLysMet 40
 Qy 121 AAATCCCTCGAGAAATCGCTATGTATATCAAGCTAACTATGAGTGCATGATAA 180
 Db 41 LysAlaSerGluLysIlePheTyValTyMetLysArgLysTyrGluAlaMetThrLys 60
 Qy 181 CTAGTTTCAGGTCACCTCCACCTTCATCGGTAGTAAAGCGGTGAGACTTCCAC 240
 Db 61 LeuGlyPheLysAlaThrLeuProProPheMetCysAsnLysArgAlaGluAspPheGln 80
 Qy 241 GGGATATTTGGTACGATCGAAACACAGGAATCAGGTGCAAGCTCTCAGATGACT 300
 Db 81 GlyAsnAspLeuAspAsnAspProAsnArgGlyAsnGlnValGluArgProGlnMetThr 100
 Qy 301 TTCCGACCTCCAGAAATCTTCCGAGATCATGCCAAGATCCAGAGCCAGAGGAGAA 360
 Db 101 PheGlyArgLeuGlnGlyIleSerProLysIleMetProLysLysProAlaGluGluGly 120
 Qy 361 AATGTTTGAGGAGTCCAGAGGATCTGGCCCAAAATATGATGGAAACAGCTGTGC 420
 Db 121 AsnAspSerGluGluValProGluAlaSerGlyProGlnAsnAspGlyLysGluLeuLys 140
 Qy 421 CCCCCGGAATCCAGTACTTGGAGAAGATTAAACAGACATCTGCACCCAAAGGGG 480
 Db 141 ProProGlyLysProThrThrSerGluLysIleHisGluArgSerGlyProLysArgGly 160
 Qy 481 AAATCCCTCGAGAAATCGCTATGTATATCAAGCTAACTATGAGTGCATGATAA 540
 Db 161 GluHisAlaThrThrHisArgLeuArgGluArgLysGlnLeuValIleTyrGluGluIle 180

Qy 541 AGCGACCTCGAGGAGATGACGAG 564
 Db 181 SerAspProGluGluAspAspGlu 188

RESULT 4
 SSX1_HUMAN
 ID SSX1_HUMAN STANDARD; PRT; 188 AA.
 AC Q16384;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE SSX1 protein (Synovial sarcoma, X breakpoint 1).
 GN SSX1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Fibrosarcoma.
 RX MEDLINE=95292974; PubMed=7539744;
 RA Crew A.J., Clark J., Fisher C., Gill S., Grimer R., Chand A.,
 RA Shipley J., Gusteron B.A., Cooper C.S.;
 RT "Fusion of SYT to two genes, SSX1 and SSX2, encoding proteins with
 RT homology to the Kruppel-associated box in human synovial sarcoma.";
 RL EMBO J. 14:2333-2340(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldi M.P., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Frange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Pahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.M., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Rutterfield Y.S.N., Krzywinski M.I., Skalska U., Smalusz D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [3]
 RP SEQUENCE OF 111-188 FROM N.A.
 RC TISSUE=Synovial sarcoma.
 RX MEDLINE=95384157; PubMed=7655467;
 RA de Leeuw B., Blawie M., Olde Weghuis D., Geurts van Kessel A.;
 RT "Identification of two alternative fusion genes, SYT-SSX1 and
 RT SYT-SSX2, in t(X;18)(p11.2;q11.2)-positive synovial sarcomas.";
 RL Hum. Mol. Genet. 4:1097-1099(1995).
 CC -!- TISSUE SPECIFICITY: Expressed at high level in the testis.
 CC -!- FUNCTION: Could act as a modulator of transcription.
 CC Expressed at low level in thyroid. Not detected in tonsil, colon,
 CC lung, spleen, prostate, kidney, striated and smooth muscles.
 CC Detected in rhabdomyosarcoma and fibrosarcoma cell lines. Not
 CC detected in mesenchymal and epithelial cell lines.
 CC -!- DISEASE: Participates in a t(X;18)(p11.2;q11.2) chromosomal
 CC translocation that is specifically found in more than 80% of
 CC synovial sarcoma and produces the SSX1-SSX1 or SSX1-SSX2 fusion
 CC products. These hybrid proteins are probably responsible for
 CC transforming activity. Heterogeneity in the position of the
 CC breakpoint can occur (low frequency).
 CC -!- SIMILARITY: Belongs to the SSX family.
 CC -!- SIMILARITY: Contains 1 KRAB-related domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X86174; AAA60110.1; ..
 DR EMBL; BC001003; AAA01003.1; ..

```
DR EMBL; S79325; AAB35378.1; -.
DR PIR; S55057; S55057.
DR Genew; HGNC:11335; SSX1.
DR MIM; 312820; -.
DR GO; GO:0005634; C:nucleus; TAS.
DR GO; GO:0003714; F:transcription co-repressor activity; TAS.
DR InterPro; IPR001909; KRAB.
DR InterPro; IPR003655; KRAB_related.
DR Pfam; PF01352; KRAB; 1.
DR SMART; SM00349; KRAB; 1.
DR PROSITE; PS50806; KRAB RELATED; 1.
DR Chromosomal translocation; Proto-oncogene, Multigene family;
KW Transcription regulation.
FT DOMAIN 20 83 KRAB-RELATED.
FT SITE 62 83 BREAKPOINT FOR TRANSLATION TO FORM THE
FT SITE 110 111 SSX1-SSX1 FUSION PROTEIN.
FT SITE 110 111 SSX1-SSX1 FUSION PROTEIN.
FT SITE 110 111 SSX1-SSX1 FUSION PROTEIN.
SQ SEQUENCE 188 AA; 21931 MW; E440D1B2AE3AE9F7 CRC64;

Alignment Scores:
Pred. No.: 8,96e-70 Length: 188
Score: 784.00 Matches: 149
Percent Similarity: 84.04% Conservative: 9
Best Local Similarity: 79.26% Mismatches: 30
Query Match: 76.49% Indels: 0
DB: 1 Gaps: 0

US-09-975-856-1 (1-576) x SSX1_HUMAN (1-188)
QY 1 ATGAAGCGAGACGACGCTTTGCAAGGAGACCCAGGATGATGCTCAAAATATCAGAGAG 60
Db 1 MetAsnGlyAspSerThrPheAlaLysArgProArgAspAlaLysAlaSerGluLys 20
QY 61 TTACGAAAGCCTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 120
Db 21 ArgSerLysAlaPheAspAlaThrTyPheSerLysLysGluLysMet 40
QY 121 AAATCTCGAGAAATCGTCTATGATGATGATGATGATGATGATGATGATGATGATGATG 180
Db 41 LysTySerGluLysLysSerTyValTyMetLysArgAsnTyLysAlaMetThrLys 60
QY 181 CTAGTTTCAAGTCACCTCCACCTTCCTGATGATGATGATGATGATGATGATGATGATG 240
Db 61 LeuGlyPheLysValThrLeuProProPheMetCysAsnLysGlnAlaThrAspPheGln 80
QY 241 GGAATGATTTTGTAACGATCGAACCACAGCAATCAGTTGATGATGATGATGATGATGAT 300
Db 81 GlyAsnAspPheAspAsnAspHisAsnArgArgIleGlnValGluHisProGlnMetThr 100
QY 301 TTCGGCAGCTCCAGAGAAATCTTCCCGAAGATCATGCCCAAGAGCCAGCAGAGAGAA 360
Db 101 PheGlyArgLeuHisArgIlelleProLysIleMetProLysLysProAlaGluAspGlu 120
QY 361 AATGTTTGAAGNAGTCCAGGATCTGCGCCCAAAATCATGGAACATGGAACATGTCG 420
Db 121 AsnAspSerLysGlyValSerGluAlaSerGlyProGlnAsnAspGlyLysGlnLeuHis 140
QY 421 CCCCAGGAAATCCAGTACCTGGAAGATTAACAGACATCTGGACCCCAAGGGGG 480
Db 141 ProGlyLysAlaAsnIleSerGluLysIleAsnLysArgSerGlyProLysArgGly 160
QY 481 AAACATCCTGGACCCACACACTGCGTGAGAGAAAGCAGCTGTTTATGAGAGATC 540
Db 161 LysHisAlaThrThrHisArgLeuArgGluArgLysGlnLeuValIleTyGluGluIle 180
QY 541 AGCAGACCTGAGGAGATGACGAG 564
Db 181 SerAspProGluGluAspAspGlu 188

RESULT 5
SSX5_HUMAN
ID SSX5_HUMAN STANDARD; PRT; 188 AA.
```

```
AC O60225; Q96AW3;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE SSX5 protein.
GN SSX5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=98021352; PubMed=9378559;
RA Gure A.O., Tuersci O., Sahin U., Tsang S., Scanlan M.J., Jager E.,
RA Knuth A., Pfeundscht M., Old L.J., Chen Y.-T.;
RT "SSX: a multigene family with several members transcribed in normal
RT testis and human cancer.";
RL Int. J. Cancer 72:965-971(1997).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RX TISSUE=Skin;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshyaki S., Carninci P., Prange C.,
RA Rana S.S., McQuellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Rulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: Could act as a modulator of transcription.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=O60225-1; Sequence=Displayed;
CC Name=2;
CC IsoId=O60225-2; Sequence=VSP_006274;
CC Note=No experimental confirmation available;
CC -!- SIMILARITY: Belongs to the SSX family.
CC -!- SIMILARITY: Contains 1 KRAB-related domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U0842; AAC05821.1; -.
DR EMBL; BC016640; AAH16640.1; -.
DR Genew; HGNC:11339; SSX5.
DR MIM; 300327; -.
DR InterPro; IPR001909; KRAB.
DR InterPro; IPR003655; KRAB_related.
DR Pfam; PF01352; KRAB; 1.
DR SMART; SM00349; KRAB; 1.
DR PROSITE; PS50806; KRAB RELATED; 1.
KW Multigene family; Transcription regulation; Alternative splicing.
FT DOMAIN 20 83 KRAB-RELATED.
FT VARSPLIC 23 23 K -> KHPWRQVCDRGHVLNLSPPVKVGRFPASSIKALLC
```

```
FT      GRGEAR (in isoform 2).
FT      /FTId=VSP_006274.
FT      P>Q (IN REF: 2).
SQ      CONFLICT 184 184
SQ      SEQUENCE 188 AA; 21628 MW; AD2A3096931CSE37 CRC64;

Alignment Scores:
Pred. No.:      2,23e-69      Length:      188
Score:          780.00        Matches:    146
Percent Similarity: 85.11%    Conservative: 14
Best Local Similarity: 77.66% Mismatches: 28
Query Match:    76.10%      Indels: 0
DB:             1           Gaps: 0

US-09-975-856-1 (1-576) x SSX5_HUMAN (1-188)
QY      1 ATCAAGCAGACGACGGCTTTCAGAGCAGACCCAGGATGATGCTCAATATACAGAGAG 60
DB      1 MetAsnGlyAspAspAlaPheValArgArgProArgValGlySerGlnIleProGlnIys 20
QY      61 TTACGAAAGCCTTCGATGATATTGCCAATATCTTCTTAAGAAAGAGTGGGAAAGATG 120
DB      21 MetGlnIysAlaPheAspAspIleAlaLysTyPheSerGlnIysGlnIrpGlnIysMet 40
QY      121 AATCTCCGAGAAATCGCTATGCTATATCAAGCTAACTATCAGCTCATGACTATAA 180
DB      41 LysAlaSerGlnIysIleIleIleIleIleIleIleIleIleIleIleIleIleIle 60
QY      181 CTAGGTTTCAAGGTCACTCCACCTTTCATCGTAGTAACAGGGCTGCAGACTTCCAC 240
DB      61 LeuGlyPheLysAlaThrLeuProPheMetArgAsnLysArgValAlaAspPheGln 80
QY      241 GGGATGATTTGGTAGCAGTCAACACACAGCAAGTCAAGTTCAGTCTCAGATGACT 300
DB      81 GlyAsnAspPheAspAsnAspProAsnArgGlyAsnGlnValGluHisProGlnMetThr 100
QY      301 TTCGCGACCTCCAGAGAAATCTCCGAGATCATGCCAAGACATGCCAAGAGCCAGCAGAGGAA 360
DB      101 PheGlyArgLeuGlnGlyIlePheProLysIleThrProGlnIysProAlaGluGluGly 120
QY      361 AATGTTTGAAGAGTCCAGAGGCATCTGGCCCAACAAATATGCGAAACAGCTGTGC 420
DB      121 AsnAspSerLysGlyValProGluAlaSerGlyProGlnAsnGlyLysGlnLeuArg 140
QY      421 CCCCCTGGGAAATCAAGTCTCTGAGAGATTAACAAGACATCTGACCCCAAGGGGG 480
DB      141 ProSerGlyLysLeuAsnThrSerGlnIysValAsnLysThrSerGlyProLysArgGly 160
QY      481 AAACATGCTGGACCCACAGACTGCTGAGAGAAAGCAGCTGTGTGTATGAGAGATC 540
DB      161 LysHisAlaTrpThrHisArgValArgGluArgLysGlnLeuValIleTyrGluGluIle 180
QY      541 AGCGACCTGAGGAGATGACGAG 564
DB      181 SerAspProGluAspAspGlu 188

RESULT 6
ZN75_HUMAN
ID      ZN75_HUMAN          STANDARD;          PRT;          289 AA.
AC      P51815;
DT      01-OCT-1996 (Rel. 34, Created)
DT      01-OCT-1996 (Rel. 34, Last sequence update)
DT      10-OCT-2003 (Rel. 42, Last annotation update)
DE      Zinc finger protein 75.
GN      ZNF75.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Lung fibroblast;
RX      MEDLINE=94116987; PubMed=8288223;
RA      Villa A., Zucchi I., Pillia G., Strina D., Susani L., Morali F.,

RA      Patroaso C., Frattini A., Lucchini F., Repetto M., Sacco M.G.,
RA      Zoppe M., Vezioni P.;
RT      "ZNF75: isolation of a cDNA clone of the KRAB zinc finger gene
RT      subfamily mapped in YACs 1 Mb telomeric of HEAT.";
RL      Genomics 18:223-229 (1993).
RN      [2]
RP      SEQUENCE OF 139-289 FROM N.A.
RA      Marino M., Archidiacono N., Franze N., Rosati M., Rocchi M.,
RA      Ballabio A., Grimaldi G.;
RL      Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
CC      -!- FUNCTION: May be involved in transcriptional regulation.
CC      -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC      -!- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-
CC      FINGER PROTEINS.
CC      -!- SIMILARITY: Contains 1 KRAB domain.
CC      -!- SIMILARITY: Contains 5 C2H2-type zinc fingers.
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC      or send an email to license@sib-sib.ch).
EMBL; S67970; AAB29696.1; -
EMBL; X68010; CAA48147.1; -
PIR; A48913; A48913.
HSSP; P25490; 1ZNM.
Genew; HGNC:13145; ZNF75.
MIM; 314997; -
GO; GO:0005634; C:nucleus; NAS.
GO; GO:0003677; F:DNA binding; NAS.
GO; GO:0008270; F:zinc ion binding; NAS.
GO; GO:0006355; P:regulation of transcription, DNA-dependent; NAS.
InterPro; IPR001909; KRAB.
InterPro; IPR007087; Znf_C2H2.
Pfam; PF01352; KRAB; 1.
Pfam; PF00096; zf_C2H2; 5.
PRINTS; PR00048; ZINCFINGER.
ProDom; PD000003; Znf_C2H2; 3.
SMART; SM00349; KRAB; 1.
SMART; SM00355; Znf_C2H2; 5.
PROSITE; PS50805; KRAB; 1.
PROSITE; PS00028; ZINC_FINGER_C2H2_1; 5.
PROSITE; PS00157; ZINC_FINGER_C2H2_2; 5.
KW      Transcription regulation; DNA-binding; Zinc-finger; Metal-binding;
KW      Nuclear protein; Repeat.
FT      DOMAIN 14 97
FT      ZN_FING 144 166
FT      ZN_FING 172 194
FT      ZN_FING 200 222
FT      ZN_FING 228 250
FT      ZN_FING 256 278
SQ      SEQUENCE 289 AA; 33683 MW; 96E7B00BF1DF64DE CRC64;

Alignment Scores:
Pred. No.:      0.00297      Length:      289
Score:          113.50        Matches:    37
Percent Similarity: 42.25%    Conservative: 23
Best Local Similarity: 26.06% Mismatches: 39
Query Match:    11.07%      Indels: 43
DB:             1           Gaps: 5

US-09-975-856-1 (1-576) x ZN75_HUMAN (1-289)
QY      73 TTCGATGATATGCCAATATCTTCTTAAGAAAGAGTGGGAAAGATCAATCTCGAG 132
DB      16 PheGluAspValAlaValTyPheSerGluGluGlnIrpGlnLeuAsnProLeuGlu 35
QY      133 AAAATCGTCTAT---GTGTATATGAAGCTAAACTATGAGTGCATCACTAACTAGGTTTC 189
DB      133 AAAATCGTCTAT---GTGTATATGAAGCTAAACTATGAGTGCATCACTAACTAGGTTTC 189
```

```

Db 36 LysThrLeuTyrrAsnAspValMetGlnAspIleTyrrGluThrValIleSerLeuGlyLeu 55
QY 190 AAGTCACTCCCTCCACCTTTTCATCGCTAGTAAACGGCTCGACAGTCTCCACGGGAATGAT 249
D 191 |||||
Db 56 LysLeuLys-----AsnAsp 60
QY 250 TTGGTACGATCGA-----AsnAsp 264
D 251 |||||
Db 61 ThrGlyAsnAspHisProIleSerValSerThrSerGluIleGlnThrSerGlyCysGlu 80
QY 265 ---AACCACAGGATCAGGTTGACGCTTCACAGATGACCTTCGCG----- 306
D 266 |||||
Db 81 ValSerLysLysThrArgMetLysIleAlaGlnLysThrMetGlyArgGluAsnProGly 100
QY 307 -----AGCCTCCAG-----AGATCTTCCGAGATCATGCCCAAGACCA 348
D 308 |||||
Db 101 AspThrHisSerValGlnLysTrpHisArgAlaGlnProArgLysLysArgLysLysPro 120
QY 349 GCAGAGGAAGAATGTTTTCAGGAGGATCCAGGATCTCGCCCAACAAATGATGGG 408
D 350 |||||
Db 121 AlaThrCysLysGlnGluLeuProLysLeuMetAspLeuHisGlyLysGlyProThrGly 140
QY 409 AACACAG 414
D 410 |||||
Db 141 GluLys 142

RESULT 7
Z398 HUMAN STANDARD; Q9P2K7; Q9UDV8; PRT; 642 AA.
AC Q9TDI7; Q9TDI8; Q9P2K7; Q9UDV8;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Zinc finger protein 398 (Zinc finger DNA binding protein p52/p71).
GN ZNF398 OR ZER6 OR KIAA1339.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RP MEDLINE=21881946; PubMed=11779858;
RX Conroy A.T., Sharma M., Holtz A.E., Wu C., Sun Z., Weigel R.J.;
RA "A novel zinc finger transcription factor with two isoforms that are
RT differentially repressed by estrogen receptor-alpha.";
RL J. Biol. Chem. 277:9326-9334 (2002).
RN [2]
SEQUENCE FROM N.A.
RP Bemis G., Langston Y., Tucci S.;
RA Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Testis;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L.H., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan B., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Sapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Haney J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Dickens M.C.,
RA Butterfield V.S.N., Krzywicki M.I., Skalska U., Smalhus D.E.,
RA Schnerch A., Schein J.F., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences."

```

```

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [4]
RP SEQUENCE OF 234-642 FROM N.A.
RX MEDLINE=20181126; PubMed=10718198;
RA Nagase T., Kikuno R., Ishikawa K.-I., Hirokawa M., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XVI.
RT The complete sequences of 150 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 7:65-73 (2000).
CC -!- FUNCTION: Function as a transcriptional activator.
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1; Synonyms=p71;
CC IsoId=Q8TDI7-1; Sequence=Displayed;
CC Name=2; Synonyms=p52;
CC IsoId=Q8TDI7-2; Sequence=VSP_006926;
CC -!- INDUCTION: By estrogen receptor alpha.
CC -!- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-
CC FINGER PROTEINS.
CC -!- SIMILARITY: Contains 1 KRAB domain.
CC -!- CAUTION: Ref.2 sequence differs from that shown due to erroneous
CC gene model prediction.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.165-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; AY049744; AAK92789.1; -
DR EMBL; AY049743; AAK92788.1; -
DR EMBL; AC004890; AAD45824.1; ALT_SEQ.
DR EMBL; EC043295; AHA43295.1; -
DR EMBL; AB037760; BAA92577.1; -
DR TRANSFAC; T05129; -
DR Genew; HGNC:18373; ZNF398.
DR GO; GO:0005634; C:nucleus; NAS.
DR GO; GO:0016563; P:transcriptional activator activity; NAS.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; NAS.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; NAS.
DR InterPro; IPR001909; KRAB_C2H2.
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF01352; KRAB; 1.
DR Pfam; PF00096; ZF-C2H2; 8.
DR ProDom; PD000003; Znf_C2H2; 2.
DR SMART; SM00349; KRAB; 1.
DR SMART; SM00355; Znf_C2H2; 8.
DR PROSITE; PS00805; KRAB; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 7.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 8.
KW Transcription regulation; Activator; DNA-binding; Zinc-finger;
KW Metal-binding; Nuclear protein; Repeat; Alternative splicing.
FT DOMAIN 143 214
FT ZN_FING 343 364 C2H2-TYPE (ATYPICAL).
FT ZN_FING 370 392 C2H2-TYPE (DEGENERATE).
FT ZN_FING 398 420 C2H2-TYPE.
FT ZN_FING 427 449 C2H2-TYPE.
FT ZN_FING 455 477 C2H2-TYPE.
FT ZN_FING 483 505 C2H2-TYPE.
FT ZN_FING 511 533 C2H2-TYPE.
FT ZN_FING 539 561 C2H2-TYPE.
FT ZN_FING 567 590 C2H2-TYPE.
FT VARSPLOC 1 171 Missing (in isoform 2).
SQ SEQUENCE 642 AA; 71311 MW; 69AA38FCD84FF633 CRC64;
/FTID=VSP_006926.

```

```

Alignment Scores:
Pred. No.: 0.0593 Length: 642
Score: 101.00 Matches: 31
Percent Similarity: 41.60% Conservative: 21
Best Local Similarity: 24.80% Mismatches: 41

```



```
Db 139 SerLeuLeuMetGluAspIlePheGlyLysGluThrProSerGlyValThrMetGluArg 158
QY 361 AATGGTTTGAAGGAA 375
Db 159 AlaGlyLeuGlyGlu 163

RESULT 9
PK3G_RAT
ID PK3G_RAT STANDARD; PRT; 1505 AA.
AC O70173;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Phosphatidylinositol-4-phosphate 3-kinase C2 domain-containing gamma
DE polypeptide (EC 2.7.1.154) [Phosphoinositide 3-Kinase-C2-gamma]
DE (PtdIns-3-kinase C2 gamma) [PI3K-C2gamma].
GN PI3K2G.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
SEQUENCE FROM N.A., FUNCTION, TISSUE SPECIFICITY, AND DEVELOPMENTAL
STAGE.
RC TISSUE=Regenerating liver;
RX MEDLINE=98184888; PubMed=9516481;
RA Ono F., Nakagawa T., Saito S., Owada Y., Sakagami H., Goto K.,
RA Suzuki M., Matsuno S., Kondo H.
RT "A novel class II phosphoinositide 3-kinase predominantly expressed in
RT the liver and its enhanced expression during liver regeneration."
RL J. Biol. Chem. 273:7731-7736(1998).
CC -!- FUNCTION: IN VITRO, PHOSPHORYLATES PTDINS AND PTDINS4P BUT NOT
CC PTDINS(4,5)P2.
CC -!- CATALYTIC ACTIVITY: ATP + 1-phosphatidyl-1D-myo-inositol 4-
CC phosphate = ADP + 1-phosphatidyl-1D-myo-inositol 3,4-bisphosphate.
CC -!- SUBCELLULAR LOCATION: Membrane-associated (By similarity).
CC -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN NORMAL LIVER. HIGH
CC LEVELS ALSO FOUND IN REGENERATING LIVER. VERY LOW LEVELS FOUND IN
CC HEART AND TESTIS.
CC -!- DEVELOPMENTAL STAGE: HIGHER LEVELS OF EXPRESSION FOUND IN ADULT
CC LIVER THAN IN FETAL LIVER.
CC -!- SIMILARITY: Belongs to the PI3/P14-kinase family.
CC -!- SIMILARITY: Contains 1 C2 domain.
CC -!- SIMILARITY: Contains 1 pbox homology (PX) domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB009636; BAA25634.1; .
CC InterPro; IPR008938; ARM.
CC InterPro; IPR000008; C2.
CC InterPro; IPR008973; C2_CaLB.
CC InterPro; IPR000403; PI3_P14_kinase.
CC InterPro; IPR002420; PI3K_C2.
CC InterPro; IPR000341; PI3K_ras_bind.
CC InterPro; IPR001263; PI3Ka.
CC InterPro; IPR001683; PX.
CC Pfam; PF00168; C2; 1.
CC Pfam; PF00454; PI3_P14_kinase; 1.
CC Pfam; PF00792; PI3K_C2; 1.
CC Pfam; PF00794; PI3K_rbd; 1.
CC Pfam; PF00613; PI3Ka; 1.
CC Pfam; PF00787; PX; 1.
CC SMART; SM00239; C2; 2.
CC SMART; SM00142; PI3K_C2; 1.
CC SMART; SM00144; PI3K_rbd; 1.
CC SMART; SM00145; PI3Ka; 1.
```

```
DR SMART; SM00146; PI3Kc; 1.
DR SMART; SM00312; PX; 1.
DR PROSITE; PS00499; C2 DOMAIN 1; FALSE NEG.
DR PROSITE; PS00004; C2 DOMAIN 2; FALSE NEG.
DR PROSITE; PS00915; PI3_4_KINASE_1; 1.
DR PROSITE; PS00916; PI3_4_KINASE_2; 1.
DR PROSITE; PS00290; PI3_4_KINASE_3; 1.
DR PROSITE; PS00195; PX; 1.
KW Transferase; Kinase; Membrane; Multigene family.
FT DOMAIN 976 1240 PI3K/PI4K.
FT DOMAIN 1259 1371 PX.
FT DOMAIN 1402 1499 C2 DOMAIN.
SQ SEQUENCE 1505 AA; 170974 MW; 5ED4C2239968C4B2 CRC64;

Alignment Scores:
Score: 0.95 Length: 1505
Pred. No.: 89.50 Matches: 33
Percent Similarity: 42.24% Conservat: 16
Best Local Similarity: 28.45% Mismatches: 50
Query Match: 8.89% Indels: 17
DB: 1 Gaps: 5

US-09-975-856-1 (1-576) x PK3G_RAT (1-1505)
QY 490 AGGCATGTTTCCCTTTTGGTCCAGATGCTTGTAAATCTTCTCAAGTACTTGAT 431
Db 83 ArgHisPheAenGluPheThrSerGln-----SerProHisPheSerGln 97
QY 430 TTCGCGGGGCGACAGCTGTTTCCCATCATTTTGTGGCCAGATGCTTGGCATTCCT 371
Db 98 LeuProPheGlyLysAlaSerAla-----IleGlyPheAenProAlaValLeuPro 114
QY 370 TCAACACATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 311
Db 115 AlaHisGlnPheIleHisGluGlyAlaSerTrpArgAenProThrArgLysTyrHisGly 134
QY 310 GCCTGCCGAAGTCACTCTGAGGACGTTCAACCTGATTCCTGTGTTTTCGATCGT---TAC 254
Db 135 GlyGluAspProArgPheSerAlaLeuThrProSerSerThrGlyLeuAspLysCysHis 154
QY 253 CAAATCATTCCTCCGGAAGTCTCAGCCGCTTCTACCGCATGAAAGTGGGAGGTGA 194
Db 155 GlnGlnGlyGlnSerGlyThrGluHisCysAsnTyrTyrVal-----Glu 169
QY 193 CTTGAAACCTAGTTTAGTACATGACCTCATGTTTAGCTTCATATACA 146
Db 170 ProGluAenAenVal-----ProHisHisTyrSerProTyrSer 182

RESULT 10
FMOL_CANFA
ID FMOL_CANFA STANDARD; PRT; 531 AA.
AC Q95LA2;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Dimethylaniline monooxygenase [N-oxide forming] 1 (EC 1.14.13.8)
DE (Hepatic flavin-containing monooxygenase 1) (FMO 1) (Dimethylaniline
DE oxidase 1).
DE FMOL.
GN FMOL.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Lattard V., Longin-Sauvageon C., Lachuer J., Burofesse T., Benoit E.;
RT "Cloning, sequencing and tissue dependent expression of FMO1 and FMO3
RT in the dog."
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: This protein is involved in the oxidative metabolism of
CC a variety of xenobiotics such as drugs and pesticides.
CC -!- CATALYTIC ACTIVITY: N,N-dimethylaniline + NADPH + O(2) = N,N-
```

dimethylalanine N-oxide + NADP(+) + H(2)O.

COFACTOR: FAD (By similarity).

SUBCELLULAR LOCATION: Mitochondrial.

TISSUE SPECIFICITY: Liver.

SIMILARITY: Belongs to the FMO family.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL: AF384053; AAK97433.1; -

InterPro: IPR000759; Adrnx_reductase.

InterPro: IPR001327; FAD_Pyr_redox.

InterPro: IPR000960; Flav_cont_mnoxgn.

PFam: PF00743; FMO-like; I.

PRINTS: PR00419; ADXRDPASE.

PRINTS: PR00368; FADPNR.

PRINTS: PR00370; FMOXYGENASE.

Oxidoreductase; Monooxygenase; NADP; Flavoprotein; FAD; Mitochondrion;

Transmembrane; Multigene family; Acetylation.

INIT MET 0 BY SIMILARITY.

MOD RES 1 1 ACETYLATION (BY SIMILARITY).

NP_BIND 8 13 FAD (ADP PART) (POTENTIAL).

NP_BIND 190 195 NADP (POTENTIAL).

SEQUENCE 531 AA; 59827 MW; 9F3458484540521F CRC64;

Alignment Scores:

Pred. No.:	0.996	Length:	531
Score:	88.50	Matches:	33
Percent Similarity:	38.21%	Conservative:	14
Best Local Similarity:	26.83%	Mismatches:	36
Query Match:	8.75%	Indels:	40
DB:	1	Gaps:	4

US-09-975-856-1 (1-576) x FMO1_CANFA (1-531)

QY 565 ACTCGTATCTCTCTCAGGTCGCTGATCTCTTCATAAACACCCAGCTGCTTTCTCTCAC 506

Db 43 ThrGluHisValGluGluGlyArgAlaSerLeuTyrIysSerVal----- 57

QY 505 GCAGTCTGTGGTCCAGGCARGTTTCCCTTTTGGTCCAGATGCTTTTAACTCTCT 446

Db 58 -----ValSerAsnSerCysLysGluMetSerCys----- 67

QY 445 CCAAGGTACTTGGATTTCCCGGGGGGCACACCTGTTCCCATCATTTTGGGCCAGATG 386

Db 68 -----TyrSerAspPhePro-----PheProGluAspTyr 77

QY 395 CCTCTGGCATTCTCTCAACACATTTCTCTCTGCTGCTTCTTGGGCATGCTCTCG 326

Db 78 ProAsnTyrValProAsnSerGluPheLeuGluTyrLeuLysMetTyrAla-Asn----- 95

QY 325 GGAAGATTCTCTGGAGGCTGCGGAAGTCATCTCAGGACGCTCAACCTGATCTCTGGT 266

Db 96 -----ArgPheSerLeuLeuLysCysII 103

QY 265 TTCGATCGTTACCAAAATCATCCCGTGGAGTCTGCAGCCCGTTTACTACGCATGAAG 206

Db 103 eArgPheLysThrLysValCysValThrLysCysProAspPheThrValThrGlyGI 123

QY 205 GTGGGAG 199

Db 123 ntrpGlu 125

RESULT 11

ID_REPL MOUSE

AC 054916; Q8C9J9; Q99LR8; PRT; 743 AA.

DT 10-OCT-2003 (Rel. 42, Created)

DT 10-OCT-2003 (Rel. 42, Last sequence update)

DE 10-OCT-2003 (Rel. 42, Last annotation update)

DE RalBP1 associated Eps domain containing protein 1 (RalBP1-interacting protein 1).

DE REP51.

GN Mus musculus (Mouse).

OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. NCBI_TaxID:10090;

RN [1] SEQUENCE FROM N.A. (ISOFORM 1), AND CHARACTERIZATION.

RC TISSUE=Muscle.

RC MEDLINE=98058900; PubMed=9395447;

RA Yamaguchi A., Urano T., Goi T., Feig L.A.;

RT "An eps homology (EH) domain protein that binds to the ral-GTPase target, RalBP1."

RL J. Biol. Chem. 272:31230-31234(1997).

RL [2]

RC SEQUENCE OF 8-743 FROM N.A. (ISOFORM 2).

RC STRAIN=C57BL/6J; TISSUE=Thymus;

RC MEDLINE=22354683; PubMed=12466851;

RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S., Nikaide I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H., Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gofjabori T., Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J., Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.M., Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S., Dalla E., Dragan T.A., Fletcher C.F., Forrest A., Frazer K.S., Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J., Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D., Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L., Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A., Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H., Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G., Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramchandran S., Ravasi T., Reed J.C., Reed D.J., Reid J., King B.Z., Ringwald M., Sadelin A., Schneider C., Semple C.A., Setou M., Shimada K., Sultana R., Takeru Y., Taylor M.S., Teasdale R.D., Tomita M., Verardo R., Wagner L., Wahlstedt C., Wang Y., Watanabe Y., Wells C., Wilming L.G., Wyshaw-Boris A., Yanagisawa M., Yang I., Yang L., Yuan Z., Zavalan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N., Hirazane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K., Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S., Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I., Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A., Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J., Birney E., Hayashizaki Y.;

RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs."

RL Nature 420:563-573(2002).

RL [3]

RN SEQUENCE OF 447-743 FROM N.A. (ISOFORM 1).

RC TISSUE=Breast tumor;

RC MEDLINE=22388257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler N.K., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J.M., Hong L., Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stappleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B., Brownstein M.J., Uddin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmitz J., Myers R.M., Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E., Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences."

```

Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[4]
RN STRUCTURE BY NMR OF 227-318.
RX MEDLINE=21285759; PubMed=11389591;
RA Kim S., Cullis D.N., Reig L.A., Balleja J.D.;
RT "Solution structure of the Rep1 EH domain and characterization of
its binding to NP7 target sequences.";
RL Biochemistry 40:6776-6785(2001).
CC -!- FUNCTION: May coordinate the cellular actions of activated EGF
receptors and Ras-GTPases.
CC -!- SUBUNIT: Homodimer (Potential). Interacts with RALBP1, CRK and
GRB2. Binding to RALBP1 does not affect its Ras-binding activity.
CC Forms a complex with the SH3 domains of CRK and GRB2 which may
link it to an EGF-responsive tyrosine kinase.
CC -!- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=O54916-1; Sequence=Displayed;
CC Name=2;
CC IsoId=O54916-2; Sequence=VSP_007956; VSP_007957;
CC Note=Due to intron retention. No experimental confirmation
available;
CC -!- TISSUE SPECIFICITY: Expressed in all tissues examined. The highest
level expression was found in the kidney and testis.
CC -!- PM: EGF stimulates phosphorylation on Tyr-residues.
CC -!- SIMILARITY: Contains 1 EH domain.
CC -!- SIMILARITY: Contains 1 EF-hand calcium-binding domain.
CC -!- CAUTION: Ref.3 sequence differs from that shown due to a
frameshift in position 719.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).
CC -----
DR EMBL; AF031939; AAB94736.1; .
DR EMBL; AK041967; BAC31117.1; ALT INIT.
DR EMBL; BC002256; AAH02256.1; ALT_FRAME.
DR PIR; T09173; T09173.
DR PDB; 1FI6; 18-JUL-01.
DR MGD; MGI:1196373; Reps1.
DR InterPro; IPR002048; EF-hand.
DR Pfam; PF00036; ehand; 1.
DR SMART; SMC0027; EH; 1.
DR PROSITE; PS00018; EF_HAND; 1.
DR PROSITE; PS00031; EH; 1.
KW Calcium-binding; Coiled coil; Phosphorylation; Alternative splicing;
3D-structure.
FT DOMAIN 233 324 EH.
FT CA_BIND 279 290 EF_HAND (POTENTIAL).
FT DOMAIN 488 551 PRO-RICH.
FT DOMAIN 599 743 INTERACTION WITH RALBP1.
FT DOMAIN 692 738 COILED COIL (POTENTIAL).
FT MOD_RES 236 236 PHOSPHORYLATION (POTENTIAL).
FT VARSPPLIC 368 402 QWETFSRSSSSQTLQFDSNIAPADPTAIVHPV ->
VSKTSLSLISLFTGRSFQDRFTAGLYQYHTP (in
isoform 2).
FT FTId=VSP_007956.
FT Missing (in isoform 2).
FT /FTId=VSP_007957.
FT VARSPPLIC 403 743
SQ SEQUENCE 743 AA; 80598 MW; 25510D11254CF4A6 CRC64;

Alignment Scores:
Pred. No.: 1.18 Length: 743
Score: 88.00 Matches: 40
Percent Similarity: 38.92% Conservatives: 25
Best Local Similarity: 23.95% Mismatches: 41
Query Match: 8.59% Indels: 61
DB: 1 Gaps: 9

```

```

US-09-975-856-1 (1-576) x REP1_MOUSE (1-743)
QY 91 TACTTCTCTAAGAAGAGTGGGAAAGATGAATCTCGGAGAAAATCGTCTATGTGTAT 150
Db 11 TyrPheGlyArgSerGlnPhe-----Tyr 18
QY 151 ATGAAGCTAAACTATGAGTTCATGATCTAACTAAGTCTTTC----- 189
Db 19 IleAlaLeuLysLeuValAlaValAlaGlnSerGlyPheProLeuArgValGluSerIle 38
QY 190 -----AAGTCCACCTCCACCTTCATGCGTACTATAA----- 222
Db 39 AsnThrValLysAspLeuProLeuProArgPheValAlaSerLysAsnGluGlnGluSer 58
QY 223 CGGGCTGCAGAGCTTCCACGGGAATGATTTTGTAAACGATCGAAACCCAC----- 270
Db 59 ArgLeuAlaAlaSerTyrSerSerAspSerGluAsnGlnGlySerTyrSerGlyValIle 78
QY 271 -----AGGAATCAGGTTCAGTCTCAGATGATCTTCGGCAGCTCCAG 315
Db 79 ProProProProGlyArgGlyGlnValLysLysGlyProGlySerHisAspAlaValGln 98
QY 316 AGAATCTTCCGAGATCATGCCCAAGAGAGCAGAGAGAGAGAAATGGTTTGAAGAA 375
Db 99 -----ProArgProSerAlaGluGln----- 106
QY 376 GTGCCAGAGCATCTGGCCCAAAATGATGCGAAACAGCTGTGCCCCCGGGAATCCA 435
Db 107 -----GluProAlaSerProValValSerProGlnGln---SerProProThrSerPro 123
QY 436 AGTACTCTGGAGAACATTAACAAGACATCTGACCCAAAGGGGGAACATG----- 487
Db 124 HisThr-TyrArgLysHisSerArgHis-----ProSerGlyLysSerGluArgPro 141
QY 488 -----CCTGCACCC 496
Db 141 OLeuThrGlyProGlyPro 147

RESULT 12
FMOL_PIG STANDARD; PRT; 531 AA.
ID AC FMOL_PIG
IC P16549;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Dimethylalanine monooxygenase [N-oxide forming] 1 (EC 1.14.13.8)
DE (Hepatic flavin-containing monooxygenase 1) (FMO 1) (Dimethylalanine
oxidase 1).
DE GN FMO1 OR FMO-1.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]_TaxID=9823;
RP SEQUENCE FROM N.A., AND SEQUENCE OF 136-150 AND 308-317.
RC TISSUE=Liver;
RX MEDLINE=90212556; PubMed=2322534;
RA Gasser R., Tynes R.E., Lawton M.P., Korsmeyer K.K., Ziegler D.M.,
RA Philpot R.M.;
RT "The flavin-containing monooxygenase expressed in pig liver: primary
sequence, distribution, and evidence for a single gene.";
RL Biochemistry 29:119-124 (1990).
RN [2]
RP SEQUENCE OF 1-13 AND 184-201, AND ACETYLATION.
RC TISSUE=Liver;
RX MEDLINE=90343821; PubMed=2383273;
RA Guan S.H., Falick A.M., Cashman J.R.;
RT "N-terminus determination: FAD and NADP binding domain mapping of hog
liver flavin-containing monooxygenase by tandem mass spectrometry.";
RL Biochem. Biophys. Res. Commun. 170:937-943 (1990).
RN [3]
RP SEQUENCE OF 185-207.

```

RC TISSUE=Liver;
 RX MEDLINE=95278229; PubMed=7758472;
 RA Wu R.-F. Ichikawa Y.;
 RT An essential lysyl residue (Lys208) in the substrate-binding site of
 RI porcine FAD-containing monooxygenase.;
 RL Eur. J. Biochem. 229:749-753(1995).
 RN [4]
 RP CARBOHYDRATE-LINKAGE SITE.
 RX MEDLINE=98451545; PubMed=9778310;
 RA Korsmeyer K.K., Guan S., Yang Z.C., Falick A.M., Ziegler D.M.,
 RA Cashman J.R.;
 RT "N-glycosylation of pig flavin-containing monooxygenase form 1:
 RI determination of the site of protein modification by mass
 RT spectrometry.";
 RL Chem. Res. Toxicol. 11:1145-1153(1998).
 CC -!- FUNCTION: This protein is involved in the oxidative metabolism of
 CC a variety of xenobiotics such as drugs and pesticides.
 CC -!- CATALYTIC ACTIVITY: N,N-dimethylaniline + NADPH + O(2) = N,N-
 CC dimethylaniline N-oxide + NADP(+) + H(2)O.
 CC -!- COFACTOR: FAD.
 CC -!- SUBCELLULAR LOCATION: Mitochondrial.
 CC -!- TISSUE SPECIFICITY: Liver.
 CC -!- SIMILARITY: Belongs to the FMO family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: M32031; AAA31033.1; -;
 DR F1R; A33768; A33768.
 DR InterPro; IPR000759; Adrnx_reductase.
 DR InterPro; IPR001327; FAD_Pyr_redox.
 DR InterPro; IPR000960; Flav_cont_mnoxgn.
 DR Pfam; PF00743; FMO-like; 1.
 DR PRINTS; PR00419; ADXEDTASE.
 DR PRINTS; PR00368; FADPNR.
 DR PRINTS; PR00370; FMOXYGENASE.
 KW Oxidoreductase; Monooxygenase.
 KW Transmembrane; Multigene family; Acetylation; Glycoprotein.
 FT INIT MET 0
 FT MOD_RES 1 0
 FT NP_BIND 8 13 ACETYLATION.
 FT NP_BIND 190 195 FAD (ADP PART) (POTENTIAL).
 FT ACT_SITE 207 207 NADP (POTENTIAL).
 FT ACT_SITE 207 207 SUBSTRATE BINDING.
 FT CARBOHYD 119 119 N-LINKED GLCNAC... (HIGH MANNOSIDE).
 SQ SEQUENCE 531 AA; 59821 MW; 58475E6D81C8157 CRC64;
 Alignment Scores:
 Pred. No.: 1.25 Length: 531
 Score: 87.50 Matches: 33
 Percent Similarity: 37.40% Conservative: 33
 Best Local Similarity: 26.83% Mismatches: 17
 Query Match: 5.69% Indels: 40
 DB: 1 Gaps: 4
 US-09-975-856-1 (1-576) x FMO1_PIG (1-531)
 QY 565 ACTCGTCTCTCTCAGGTCGTCATCTCTTCATTAACACACGAGTCGTTCTCTCAC 506
 Db 43 ThrGluHisValGluGluGlyArgAlaSerLeuTyrLysSerVal----- 57
 QY 505 GCAGTCTGGTCCAGGCAGTCTTCCCTTTTGGGTCCAGAGTCTTGTATATCTTCT 446
 Db 58 -----ValSerAsnSerCysLysGluMetSerCys-----° 67
 QY 445 CCAAGGTACTGGATTTCCTGGGGGGGACAGCTGTTTCCCATCATTTTGGGGCAGATG 386
 Db 68 -----TyrProAspPhePro-----PheProGluAspTyr 77

QY 385 CCTCTGGCACTTCTCTCAACACATTTCTCTCTGCTGGCTTCTTGGGCATGATCTTCG 326
 Db 78 ProAntyrrValProAsnSerHisPheLeuGluTyrLeu----- 90
 QY 325 GGAAGATTCTCTGGAGGCTGCCGAAAGTCATCTGAGGACGTCCTAACCTGATCTCTGGT 266
 Db 91 -----Arg-MetTyrAlaAsnGlnPheAsnLeuLeuLysCysI 103
 QY 265 TTCGATCGTTACCAAAATCATTCCGCTGGAAGTCTGCAGCCGCTTTACTAGCATGAAG 206
 Db 103 eGlnPheLysThrLysValCysSerValThrLysHisGluAspPheAsnThrThrGlyG 123
 QY 205 GTGGGAG 199
 Db 123 nTrpAsp 125
 RESULT 13
 REPI_HUMAN
 ID REPI_HUMAN STANDARD; PRT; 744 AA.
 AC Q96D71; QBNDR7; Q8WU62; Q9BX99;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE RalBP1 associated Eps domain containing protein 1 (RalBP1-interacting
 DE protein 1).
 GN REPS1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RA Mao Y., Xie Y., Zhou Z., Zhao W., Zhao S., Wang W., Huang Y., Wang S.,
 RA Tang R., Chen X., Wu C.;
 RL Submitted (Mar-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORMS 2 AND 3).
 RC TISSUE=Brain, and Lymph;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Rana S.S., Lofellano N.A., Peters G.J., Abramson R.D., Mullen S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley D.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Prange C.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [3]
 RP SEQUENCE OF 274-743 FROM N.A. (ISOFORM 1).
 RC TISSUE=Brain;
 RA Bioecker H., Boecher M., Brandt P., Mewes H.-W., Weil B., Wiemann S.;
 RL Submitted (Jul-2002) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: May coordinate the cellular actions of activated EGF
 CC receptors and Ral-GTPases (By similarity).
 CC -!- SUBUNIT: Homodimer (potential). Interacts with RALBP1, CRK and
 CC GRB2. Binding to RALBP1 does not affect its Ral-binding activity.
 CC Forms a complex with the SH3 domains of CRK and GRB2 which may
 CC link it to an EGF-responsive tyrosine kinase (By similarity).
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event-Alternative splicing; Named isoforms=3;
 CC Name=1;

```

CC      IsoId=Q96D71-1; Sequence=Displayed;
CC      Name=2;
CC      IsoId=Q96D71-2; Sequence=VSP_007953, VSP_007954;
CC      Name=3;
CC      IsoId=Q96D71-3; Sequence=VSP_007955;
CC      -!- PPM: EGF stimulates phosphorylation on Tyr-residues (By
CC      similarity);
CC      -!- SIMILARITY: Contains 1 EH domain.
CC      -!- SIMILARITY: Contains 1 EF-hand calcium-binding domain.
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; AF251052; AAK34942.1; -
DR      EMBL; BC012764; AAH12764.1; -
DR      EMBL; BC021211; AAH21211.1; -
DR      EMBL; AL331900; CAD38569.1; -
DR      Genew; HGNC:15578; REPS1.
DR      InterPro; IPR002048; EF-hand.
DR      InterPro; IPR002048; EFS15_homology.
DR      SMART; SM00027; EH; 1.
DR      SMART; SM00018; EF_HAND; 1.
DR      PROSITE; PS00031; EH; 1.
DR      Calcium-binding; Coiled coil; Phosphorylation; Alternative splicing.
KW      DOMAIN 233 324
FT      CA BIND 279 290
FT      DOMAIN 489 552
FT      DOMAIN 600 744
FT      DOMAIN 693 739
FT      MOD RES 236 236
FT      VARSPIC 368 394
FT      VARSPIC 458 521
FT      VARSPIC 458 458
FT      VARSPIC 576 576
FT      CONFLICT 665 665
FT      CONFLICT 742 743
SQ      SEQUENCE 744 AA; 80769 MW; 1DFF29711DB2B5B4 CRC64;

Alignment Scores:
Pred. No.: 1.49 Length: 744
Score: 87.00 Matches: 39
Percent Similarity: 38.89% Conservative: 24
Best Local Similarity: 24.07% Mismatches: 48
Query Match: 8.49% Indels: 51
DB: 1 Gaps: 9

US-09-975-856-1 (1-576) x REP1_HUMAN (1-744)
Qy 91 TACTTCTTAAGAAGAGTGGGAAGATGAATCCTCGGAGAAAATCGTCTATGCTAT 150
Db 11 TyrPheGlyArgSerGlnPhe-
Qy 151 ATGAAGCTAACTATGAGTCTACGTAACTAGGTTTC----- 189
Db 19 IleAlaLeuLysLeuValAlaValAlaGlnSerGlyPheProLeuArgValGluSerIle 38
Qy 190 -----AAGGTACCTCCCTCCACTTTCAGTGTAGTAA----- 222
Db 39 AsnThrValLysAspLeuProLeuProArgPheValAlaSerLysAsnGluGlnGluSer 58
Qy 223 CGGCTGCAGACTCCCGGGAATGATTTTGGTAACGATCGAACCACAGGAATCAGGTT 282
Db 59 ArgHisAlaAlaSerTyrSerSerAspSerGluAsnGlnGlnGlySerTyrSerGlyValIle 78

```

```

Qy 283 GAACGTCTCT-----CAGATGACTTTTCGGCGCCTCCAGAGAAATCTTCCCGAAG 330
Db 79 PropProProGlyArgGlyGlnValLysLysGlySerValSerHis-----AspThr 96
Qy 331 ATCATGCCCAAGAACCCAGACAGAGAGAAAATGGTTTGAAGGAAGTCCAGAGGCATCT 390
Db 97 ValGlnProArgThrSerAlaAspAlaGln-----GluProAla 109
Qy 391 GGCACCAAAAATGATGGGAAACAGCTGTGCCCCCGGGAATCCAAAGTACCTTGAGAG 450
Db 110 SerProValValSerProGlnGln--SerProProThrSerProHisThr-TyrArgLys 128
Qy 451 ATTAACAAGACATCTGGACCCCAAAAGGGGGAACATG-----CCTGG 492
Db 128 HisSerArgHis-----ProSerGlyGlyAsnSerGluArgProLeuAlaGlyProGln 146
Qy 493 ACCC 496
Db 146 yPro 147

RESULT 14
SK11_CHICK
ID SK11_CHICK STANDARD; PRT; 396 AA.
AC P48435;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Transcription factor SOX-11.
GN SOX11.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]_TaxID=9031;
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=95267693; PubMed=7748786;
RA Uwanogho D., Rex M., Cartwright E.J., Pearl G., Healy C.,
RA Scotting P.J., Sharpe P.T.;
RT "Embryonic expression of the chicken Sox2, Sox3 and Sox11 genes
RT suggests an interactive role in neuronal development.";
RL Mech. Dev. 49:23-36(1995).
CC -!- FUNCTION: May function as switches in neuronal development.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- TISSUE SPECIFICITY: Low level expression is seen in
CC undifferentiated proliferating cells of neural epithelium. A
CC greater expression is seen in the maturing neurons after they
CC leave the neural epithelium. It is also found in the gut
CC epithelium and adrenal medulla.
CC -!- SIMILARITY: Contains 1 HMG box domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR      EMBL; U12534; AAB09664.1; -
DR      PIR; I50707; I50707.
DR      HSSP; P48436; 1SX9.
DR      InterPro; IPR000910; HMG_12_box.
DR      Pfam; PF00505; HMG_box; 1.
DR      SMART; SM00398; HMG; 1.
DR      PROSITE; PS0118; HMG_BOX_2; 1.
KW      Transcription regulation; DNA-binding; Nuclear protein.
FT      DNA BIND 49 117
FT      DOMAIN 162 165
FT      POLY-ALA 162 165
FT      DOMAIN 284 294
FT      POLY-GLU 284 294
FT      POLY-PRO 288 294
FT      POLY-ARG 332 335

```


DR EMBL; X07290; CAA30269.1; -.
 DR PIR; S00754; S00754.
 DR Genew; HGNC:13104; ZNF38.
 DR Genew; HGNC:13089; ZNF3.
 DR MIM; 601261; -.
 DR MIM; 194510; -.
 DR GO; GO:0005634; Cnucleus; IC.
 DR GO; GO:0003700; F:transcription factor activity; NAS.
 DR GO; GO:0008270; F:zinc ion binding; NAS.
 DR GO; GO:0045321; P:cell activation; NAS.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; NAS.
 DR InterPro; IPR001909; KRAB.
 DR InterPro; IPR007087; Znf_C2H2.
 DR InterPro; IPR007086; Znf_C2H2_sub.
 DR Pfam; PF01352; KRAB; 1.
 DR PRINTS; PR00096; zf-C2H2; 8.
 DR ProDom; PD000003; Znf_C2H2; 8.
 DR SMART; SM00349; KRAB; 1.
 DR SMART; SM00355; Znf_C2H2; 8.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 8.
 DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 8.
 DR PROSITE; PS0805; KRAB; 1.
 KW Transcription regulation; Activator; Zinc-finger; DNA-binding;
 KW Repeat; Metal-binding; Nuclear protein; Differentiation.
 FT DOMAIN 51 123
 FT ZN_FING 200 227
 FT ZN_FING 228 255
 FT ZN_FING 256 283
 FT ZN_FING 284 311
 FT ZN_FING 312 339
 FT ZN_FING 340 367
 FT ZN_FING 368 395
 FT ZN_FING 396 423
 FT CONFLICT 21 21 L->P (IN REF. 2).
 FT CONFLICT 252 256 GERPY -> IRSG (IN REF. 5).
 FT CONFLICT 336 387 GERPYECNEGKAFSSSHLYQHRIHTGERPYECMBCGK
 FT FTYSGLIQQH -> EALFTVTLRLPLSDPIVTNEAAF
 FT PAESLATIFALIWRLFVHSLMPKKV (IN REF. 3).
 SQ SEQUENCE 446 AA; 50932 MW; 67A5926807304782 CRC64;

Alignment Scores:

Pred. No.:	1.53	Length:	446
Score:	86.50	Matches:	42
Percent Similarity:	36.98%	Conservative:	29
Best Local Similarity:	21.88%	Mismatches:	50
Query Match:	8.44%	Indels:	71
DB:	1	Gaps:	8

US-09-975-856-1 (1-576) x ZN38_HUMAN (1-446)

QY	40	GATGCTCAATATACAGAGAAGTTACGAAAGGCC-----TTGGAT	78
Db	35	AspGluMetLeuAlaAlaLeuLeuLysSerGlnGluValThrPheGlu	54
QY	79	GATATTGCCAATACCTCTTAAGAAGAGTGGAAAAAGATGAATCCTCGGAGAAATC	138
Db	55	AspValAlaValPyrPheIleArgLysGluTrpLysArgLeuGluProIaGlnArgAsp	74
QY	139	GTCTATGTGTATGAAGCTA---AACTATGAGGTCTATGACTAAACTAGGTTTCAAGGTC	195
Db	75	LeuTyrArgAspValMetLeuGluAsnTyr-----	84
QY	196	ACCTCCACCTTTTCATGCGTAGCTAAACGGGCTGCAGACTTCACGGGAATGATTTGGT	255
Db	85	-----GlyAsnValPheSer	89
QY	256	AACGATCGAAACACACAGG-----AATCAGGTTGAACGTCCTCAGATGACTTTC	303
Db	90	LeuAspArgGluThrArgThrGluAsnAspGlnGluIleSerGluAspThrArgSerHis	109
QY	304	GGCAGCCTCCAGAGAATCTTCCGAAG-----ATCATGCCCAAGAGCA	348

Db	110	GlyValLeuLeuGlyArgPheGlnLysAspIleSerGlnGlyLeuLysPheLysGluAla	129
QY	349	GCAGAGGAAGAAAATGTTTCAAGGAAGTCCAGAGGCATCTGCCCAAAAATGATGGG	408
Db	130	TyrGluArgGluValSerLeuLysArg-----	138
QY	409	AAACAGCTGTGCCCCCGGGAATCCAAATCCAACTTGGAGAAGATTAAACAAGACATCTGGA	468
Db	139	-----ProLeuGlyAsn---SerProGlyGluArgLeuAsnArg-----	150
QY	469	CCCAAAAGGGGAAACATGCTGTGAGCCACAGACTGCGTGAGAGAAAAGCAGCTGGTGGTT	528
Db	151	-----LysMetProAspPheGlyGlnValThrVal	160
QY	529	TATGAAGAGATACCGACCCCTGAGGAAGATGACGAG	564
Db	161	GluGluLysLeuThrProArgGlyGluArgSerGlu	172

Search completed: March 31, 2004, 13:58:04
 Job time : 18 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: March 31, 2004, 13:40:37 ; Search time 47 Seconds
(without alignments)
7733.553 Million cell updates/sec

Title: US-09-975-856-1

Perfect score: 1025

Sequence: 1 ATGACGAGGACGACGCCTT.....ATGACGAGTAACCTCCCTCG 576

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Zgapop 6.0 , Zgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 2034082

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame_n2p.model -DEV=xlp
-Q/csn2_1/USFTO_spool_3/US09975856/runat_31032004_132903_14748/app_query.fasta_1.775
-DB=SPREMBL_25 -QFMT=fastaan -SUFFIX=rspt -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09975856 @CGN 1.1.86 @runat_31032004_132903_14748 -NCPU=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

SPREMBL_25: *
1: sp_archaea: *
2: sp_bacteria: *
3: sp_fungi: *
4: sp_human: *
5: sp_invertebrate: *
6: sp_mammal: *
7: sp_mhc: *
8: sp_organelle: *
9: sp_phase: *
10: sp_plant: *
11: sp_rodent: *
12: sp_virus: *
13: sp_vertebrate: *
14: sp_unclassified: *
15: sp_rvirus: *
16: sp_bacteria: *
17: sp_archae: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	685.5	66.9	223	4	Q9BU88 homo sapien

2	636	62.0	170	4	Q9BRW7	Q9brw7 homo sapien
3	576	56.2	155	4	Q96Q11	Q96q11 homo sapien
4	545.5	53.2	136	4	Q96Q10	Q96q10 homo sapien
5	363	35.4	98	4	Q9Y444	Q9y444 homo sapien
6	255	24.9	64	4	Q8WZ29	Q8wz29 homo sapien
7	247.5	24.1	113	4	Q9NZK4	Q9nzk4 homo sapien
8	214.5	20.9	128	11	Q8OZT4	Q8ozt4 mus musculus
9	214.5	20.9	165	11	Q8C5Z3	Q8c5z3 mus musculus
10	211.5	20.6	170	11	Q9CPU1	Q9cpu1 mus musculus
11	160.5	15.7	117	4	Q75101	Q75101 homo sapien
12	137	13.4	39	4	Q81ZHO	Q81zh0 homo sapien
13	137	13.4	39	4	Q81ZG9	Q81zg9 homo sapien
14	137	13.4	39	4	Q81ZG8	Q81zg8 homo sapien
15	137	13.4	39	4	Q81ZG7	Q81zg7 homo sapien
16	137	13.4	39	4	Q81ZG6	Q81zg6 homo sapien
17	137	13.4	39	4	Q81ZG5	Q81zg5 homo sapien
18	124	12.1	39	4	Q81ZH1	Q81zh1 homo sapien
19	113.5	11.1	510	4	Q86TD5	Q86td5 homo sapien
20	112.5	11.0	1111	4	Q60290	Q60290 homo sapien
21	111	10.8	330	4	Q96XN0	Q96xn0 homo sapien
22	110.5	10.8	524	4	Q8N2J5	Q8n2j5 homo sapien
23	110.5	10.8	785	4	Q9ULD5	Q9uld5 homo sapien
24	109.5	10.7	652	11	Q8BFS8	Q8bfs8 mus musculus
25	109	10.6	432	3	Q9HF87	Q9hfs7 botrytis ci
26	109	10.6	635	11	Q8OZZ2	Q8ozz2 mus musculus
27	106.5	10.4	782	4	Q8N393	Q8n393 homo sapien
28	106	10.3	579	11	Q9WV10	Q9wv10 mus musculus
29	104.5	10.2	405	4	Q81W91	Q81w91 homo sapien
30	104.5	10.2	579	11	Q8C393	Q8c393 mus musculus
31	104	10.1	643	11	Q8BV16	Q8bv16 mus musculus
32	103	10.0	614	11	Q8C887	Q8c887 mus musculus
33	103	10.0	614	11	Q8BPJ3	Q8bpj3 mus musculus
34	103	10.0	621	11	Q8OTC5	Q8otc5 mus musculus
35	102	10.0	511	11	Q8BVH0	Q8bvh0 mus musculus
36	102	10.0	649	4	Q9P215	Q9p215 homo sapien
37	100.5	9.8	182	4	Q9UDV5	Q9udv5 homo sapien
38	100.5	9.8	431	11	Q8C964	Q8c964 mus musculus
39	98.5	9.6	258	17	Q8TPS6	Q8tps6 methanosarc
40	98.5	9.6	428	17	Q8TUH2	Q8tuh2 methanosarc
41	98.5	9.6	428	17	Q8TYT7	Q8tyt7 methanosarc
42	98.5	9.6	428	17	Q8TNA0	Q8tna0 methanosarc
43	98.5	9.6	428	17	Q8TMC5	Q8tmc5 methanosarc
44	98.5	9.6	428	17	Q8TL32	Q8tl32 methanosarc
45	98.5	9.6	428	17	Q8TIL4	Q8til4 methanosarc

ALIGNMENTS

RESULT 1

Q9BU88	PRELIMINARY;	PRT;	223 AA.
ID Q9BU88			
AC Q9BU88			
DT 01-JUN-2001 (TREMBLrel. 17, Created)			
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)			
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)			
DE Similar to synovial sarcoma, X breakpoint 2.			
OS Homo sapiens (Human)			
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX NCBI_TaxID=9606;			
RN [1]			
RP SEQUENCE FROM N.A.			
RC TISSUE=Placenta;			
RA Strausberg R.;			
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.			
DR EMBL; BC02818; AA02818.1; -			
DR GO; GO:0005622; C:intracellular; IEA.			
DR GO; GO:0003676; F:nucleic acid binding; IEA.			
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.			
DR InterPro; IPR001909; KRAB.			
DR InterPro; IPR003655; KRAB_related.			
DR Pfam; PF01352; KRAB; 1.			
DR SMART; SM00349; KRAB; 1.			

DR PROSITE; PS50806; KRAB RELATED; 1.
SQ SEQUENCE 223 AA; 25173 MW; 2BF8B1FFA4D58094 CRC64;
Alignment Scores:
Pred. No.: 1.15e-62 Length: 223
Score: 685.50 Matches: 137
Percent Similarity: 66.07% Conservative: 11
Best Local Similarity: 61.16% Mismatches: 26
Query Match: 66.88% Indels: 50
DB: 4 Gaps: 1

US-09-975-856-1 (1-576) x Q9BU88 (1-223)

QY 1 ATGAACGAGACGACGCTTTGCAAGGAGACCCAGGATGATCTCAAAATATCAGAGAAG 60
DB 1 MetAsnGlyAspAlaPheAlaArgProThrValGlyAlaGlnIleProGluLys 20
QY 61 TTACGAAAGCCCTTCGATGATATTGCCAATATCTCTAAGAAAGAGTGGGAAGAGT 120
DB 21 IleglnLysAlaPheAspAlaLysPheSerLysGluGlnTrpGluLysMet 40
QY 121 AAATCCTCGAGAAATCGCTATGTATATGAAGCTAACTATCAGGTCATGACTAAA 180
DB 41 LysAlaSerGluLysIlePheTyValTyMetLysArgLysTyGluAlaMetThrLys 60
QY 181 CTAGGTTTCAAGTCACCTCCACCTTCATCGTAGTAAGCGGCTCCAGCTCCAC 240
DB 61 LeuGlyPheLysAlaThrLeuProPheMetCysAsnLysArgAlaGluAspPheGln 80
QY 241 GGAATGATTTCGTAAACGATCGAAACACACAGGAATCAGGTTCAACGCTCCTCAGATGACT 300
DB 81 GlyAsnAspLeuAspAsnAspProAsnArgGlyAsnGlnValGluArgProGlnMetThr 100
QY 301 TTCCGACGCTCCAGAGAAATCTCCCGAGATCATGCCAGAGCCAGACAGAGAGAA 360
DB 101 PheGlyArgLeuGlnValProGluAlaSerGlyProGlnAsnAspGlyLysGluLeuCys 120
QY 361 AATGGTTTGAAGAAAGTCCAGAGGATCTGCCACAAATATGCGAAACAGCTGTGC 420
DB 421 CCCCCGGGAATCCAGTACTCTGGAGAGATTAAACAGACATCTCGA 468
QY 141 ProGlyLysProThrThrSerGluLysIleHisGluArgSerGlyAsnArgGluAla 160
QY 468 ----- 468
DB 161 GlnGluLysGluGluArgArgGlyThrAlaHisArgTrpSerSerGlnAsnThrHisAsn 180
QY 468 ----- 468
DB 181 IleGlyArgPheSerLeuSerThrSerMetGlyAlaValHisGlyThrProLysThrIle 200
QY 469 -----CCCAAAGGGGAACATGCTGGACCCACAGACTGGTGAGAGA 513
DB 201 ThrHisAsnArgAspPro-LysGlyGlyAsnMetProGlyProThrAspCysValArgG 220
QY 514 AAGCAGCTGG 523
DB 220 uAsnSerTrp 223

RESULT 2

Q9BRW7 PRELIMINARY; PRT; 170 AA.

AC Q9BRW7;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Synovial sarcoma, X breakpoint 3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RP SEQUENCE FROM N.A.

RP SEQUENCE FROM N.A.
RC TISSUE=Bone marrow;
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC005904; AA05904.1; -.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0003676; P:nucleic acid binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001909; KRAB.
DR InterPro; IPR003655; KRAB-related.
DR Pfam; PF01352; KRAB; 1.
DR SMART; SM00349; KRAB; 1.
DR PROSITE; PS50806; KRAB RELATED; 1.
SQ SEQUENCE 170 AA; 19457 MW; D981F807A9C7EAB1 CRC64;

Alignment Scores:

Pred. No.: 1.61e-57 Length: 170
Score: 636.00 Matches: 122
Percent Similarity: 82.05% Conservative: 6
Best Local Similarity: 78.21% Mismatches: 28
Query Match: 62.05% Indels: 0
DB: 4 Gaps: 0

US-09-975-856-1 (1-576) x Q9BRW7 (1-170)

QY 1 ATGAACGAGACGACGCTTTGCAAGGAGACCCAGGATGATCTCAAAATATCAGAGAAG 60
DB 1 MetAsnGlyAspAlaPheAlaArgProThrValGlyAlaGlnIleProGluLys 20
QY 61 TTACGAAAGCCCTTCGATGATATTGCCAATATCTCTAAGAAAGAGTGGGAAGAGT 120
DB 21 IleglnLysAlaPheAspAlaLysPheSerLysGluGlnTrpGluLysMet 40
QY 121 AAATCCTCGGAGAAATCGCTATGTATATGAAGCTAACTATGAGTCAAGTCACTAAA 180
DB 41 LysValSerGluLysIleValTyValTyMetLysArgLysTyGluAlaMetThrLys 60
QY 181 CTAGGTTTCAAGTCACCTCCACCTTCATCGTAGTAAGCGGCTCCAGCTCCAC 240
DB 61 LeuGlyPheLysAlaIleLeuProSerPheMetArgAsnLysArgValThrAspPheGln 80
QY 241 GGAATGATTTCGTAAACGATCGAAACACACAGGAATCAGGTTCAACGCTCCTCAGATGACT 300
DB 81 GlyAsnAspPheAspAsnAspProAsnArgGlyAsnGlnValGlnArgProGlnMetThr 100
QY 301 TTCCGACGCTCCAGAGAAATCTCCCGAGATCATGCCAGAGCCAGACAGAGAGAA 360
DB 101 PheGlyArgLeuGlnGlyIlePheProLysIleMetProLysLysProAlaGluGluGly 120
QY 361 AATGGTTTGAAGAAAGTCCAGAGGATCTGCCACAAATATGATGGAACAGCTGTGC 420
DB 121 AsnValSerLysGluValProGluAlaSerGlyProGlnAsnAspGlyLysGlnLeuCys 140
QY 421 CCCCCGGGAATCCAGTACTCTGGAGAGATTAAACAGACATCTGGA 468
DB 141 ProGlyLysProThrThrSerGluLysIleAsnMetIleSerGly 156

RESULT 3

Q9EQ11 PRELIMINARY; PRT; 155 AA.

AC Q9EQ11;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE DJ54B20.1.1 (Novel SSX family protein (isoform 1)) (Fragment).
GN DJ54B20.1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RP SEQUENCE FROM N.A.

RA Grafham D.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: Z98304; CAC41946.1; -.
DR GO: GO:0005622; C:intracellular; IEA.
DR GO: GO:0003676; F:nucleic acid binding; IEA.
DR GO: GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO: GO:0003676; F:nucleic acid binding; IEA.
DR GO: GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro: IPR001909; KRAB.
DR InterPro: IPR003655; KRAB_related.
DR Pfam: PF01352; KRAB; 1.
DR SMART: SM00349; KRAB; 1.
DR PROSITE: PS0806; KRAB_RELATED; 1.
FT NON_TER 155
SQ SEQUENCE 155 AA; 17689 MW; 06EC7C9D935A1A30 CRC64;
Alignment Scores:
Pred. No.: 2,9e-51 Length: 155
Score: 576.00 Matches: 113
Percent Similarity: 80.65% Conservative: 12
Best Local Similarity: 72.90% Mismatches: 30
Query Match: 56.20% Indels: 0
DB: 4 Gaps: 0
US-09-975-856-1 (1-576) x Q96Q10 (1-155)
QY 1 ATGAACGGAGACGCGCTTTGCAAGGAGACCCAGGATGATGCTCAATATCAGAGAAG 60
Db 1 MetAsnGlyAspAlaPheAlaLysArgProArgAspAlaLysArgLys 20
QY 61 TTACGAAAGCCTTCGATATGTCCTTCTTAAGAAAGAGTGGGAAAGATG 120
Db 21 ArgSerLysAlaPheAspAlaLysArgProArgAspAlaLysArgLysMet 40
QY 121 AAATCTCGAGAAATCGTCTATGTCATGATCAAGCTAACTATGAGTCATGATAA 180
Db 41 LysPheSerGluLysIleSerCysValHisMetLysArgLysTyrGluAlaMetThrLys 60
QY 181 CTAGGTTTCAAGGTCACCTCCACCTTCATGCGTAGTAAAGCGGTCGAGACTTCCAC 240
Db 61 LeuGlyPheAsnValThrLeuSerLeuPheMetArgAsnLysArgAlaThrAspSerGln 80
QY 241 GGGATATGTTGGTACGATCGAACCACAGGATCAGTGAAGTCTCAGATGACT 300
Db 81 ArgAsnAspSerAspAsnAspArgAsnArgGlyAsnGluValGluArgProGlnMetThr 100
QY 301 TTCCGACGCTCCAGAGATCTTCCCGAAGATCATGCCAAGAGCCAGCAGAGGAGAA 360
Db 101 PheGlyArgLeuGlnArgIleProLysIleMetProGluLysProAlaGluGluGly 120
QY 361 AATGGTTTGAAGAGTCCAGAGGATCTGCGCCCAAAATGATGGGAAACAGCTGTC 420
Db 121 SerAspSerLysGlyValProGluAlaSerGlyProGlnAsnAspGlyLysLysLeuCys 140
QY 421 CCCCCGGAAATCCCAAGTACCTTGGAGAAGATTAAACAGACATCT 465
Db 141 ProProGlyLysAlaSerSerSerGluLysIleHisGluArgSer 155
RESULT 4
Q96Q10 PRELIMINARY; PRT; 156 AA.
AC Q96Q10;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE DJS4B20.1.2 (Novel SSX family protein (isoform 2)) (Fragment).
GN DJS4B20.1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN 1;
RP SEQUENCE FROM N.A.
RA Grafham D.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL: Z98304; CAC41947.1; -.
DR GO: GO:0005622; C:intracellular; IEA.
DR GO: GO:0003676; F:nucleic acid binding; IEA.
DR GO: GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro: IPR001909; KRAB.
DR InterPro: IPR003655; KRAB_related.
DR Pfam: PF01352; KRAB; 1.
DR SMART: SM00349; KRAB; 1.
DR PROSITE: PS0806; KRAB_RELATED; 1.
FT NON_TER 196
SQ SEQUENCE 196 AA; 22299 MW; 71CALBE1F4BA07F8 CRC64;
Alignment Scores:
Pred. No.: 4,59e-48 Length: 196
Score: 545.50 Matches: 113
Percent Similarity: 63.78% Conservative: 12
Best Local Similarity: 57.65% Mismatches: 30
Query Match: 53.22% Indels: 41
DB: 4 Gaps: 1
US-09-975-856-1 (1-576) x Q96Q10 (1-196)
QY 1 ATGAACGGAGACGCGCTTTGCAAGGAGACCCAGGATGATGCTCAATATCAGAGAAG 60
Db 1 MetAsnGlyAspAlaPheAlaLysArgProArgAspAlaLysArgLys 20
QY 61 TTACGAAAGCCTTCGATATGTCCTTCTTAAGAAAGAGTGGGAAAGATG 120
Db 21 ArgSerLysHisProTyrArgLysValCysAspLeuAlaLeuHisLeuValThrLeuThr 40
QY 69 ----- 69
Db 41 ProPheTyrLysValGlyArgGluProAlaSerIleThrGluAlaLeuLeuCysGlyArg 60
QY 70 ----- GCCTTCGATGATATTCGCAAAATCTCTCTAAGAAAGAGTGGGAAAG 117
Db 61 GlyGluAlaArgAlaPheAspAlaLysArgProArgAspAlaLysArgLys 80
QY 118 ATGAATCTCGAGAAATCGTCTATGTCATGATGATGATGATGATGATGATGATGATGAT 177
Db 81 MetLysPheSerGluLysIleSerCysValHisMetLysArgLysTyrGluAlaMetThr 100
QY 178 AAATGAGTTTCAAGTCCACCTCCACCTTCATGCGTAGTAAAGCGGTCGAGACTTC 237
Db 101 LysLeuGlyPheAsnValThrLeuSerLeuPheMetArgAsnLysArgAlaThrAspSer 120
QY 238 CACGGAGATGATTTGGTAAAGTCAAGATCGAACCACAGGATCAGGTTGACCTCCAGATG 297
Db 121 GlnArgAsnAspSerAspAsnAspArgAsnArgGlyAsnGluValGluArgProGlnMet 140
QY 298 ACTTTCCGACGCTCCAGAGATCTTCCCGAAGATCATGCCAAGAGCCAGCAGAGGAA 357
Db 141 ThrPheGlyArgLeuGlnArgIleProLysIleMetProGluLysProAlaGluGlu 160
QY 358 GAAATGTTTGAAGAGTCCAGAGGATCTGCGCCCAAAATGATGGGAAACAGCTG 417
Db 161 GlySerAspSerLysGlyValProGluAlaSerGlyProGlnAsnAspGlyLysLysLeu 180
QY 418 TCCCCCGGGAAATCCCAAGTACCTTGGAGAAGATTAAACAGACATCT 465
Db 181 CysProGlyLysAlaSerSerSerGluLysIleHisGluArgSer 196
RESULT 5
Q9Y444 PRELIMINARY; PRT; 98 AA.
AC Q9Y444;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE SYT-SSX protein (fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```
OK NCBI_TaxID=9606;
RP SEQUENCE FROM N.A.
RA CLARK J., Rocques P.J., Crew A.J., Gill S., Shipley J., Chan A.M.,
RX MEDLINE=95038836; PubMed=7951320;
RA Clark J., Rocques P.J., Crew A.J., Gill S., Shipley J., Chan A.M.,
RT "Identification of a novel gene, SYT and SSX, involved in the t(X;18)
RT (p11.2;q11.2) translocation found in human synovial sarcoma.";
RL Nat. Genet. 7:502-508(1994).
DR EMBL; X79200; CAB36970.1; -.
FT NON TER 1
SQ SEQUENCE 98 AA; 11136 MW; 1C71F5C8D54513DD CRC64;

Alignment Scores:
Pred. No.: 4,56e-29 Length: 98
Score: 363.00 Matches: 65
Percent Similarity: 91.36% Conservative: 9
Best Local Similarity: 80.25% Mismatches: 7
Query Match: 35.41% Indels: 0
DB: 4 Gaps: 0

US-09-975-856-1 (1-576) x Q9Y444 (1-98)
QY 322 TTCGCGAAGATCATGCCAGAGAGCCAGAGAGAGAAATGTTTGAAGAGAGTCCCA 381
Dy :: ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Dy 18 TyrAspGlnMetProLysLysProAlaGluGlyAsnAspSerGluGluValPro 37
QY 382 GAGCATCTGCGCCACACAAATGATGGGAAACAGCTGTGCCCGGGAAATCCCAAGTAC 441
Dy ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Dy 38 GluAlaSerGlyProGlnAsnAspGlyGlyGluLeuCysProGlyLysProThrThr 57
QY 442 TTGAGAGAGATTAAACAGACATCTGGACCCCAAGGGGAAACATGCTCGACCCACAGA 501
Dy ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Dy 58 SerGluLysIleHisGluArgSerGlyProLysArgGlyGluHisAlaTyrThrHisArg 77
QY 502 CTGCGTGAAGAGAGAGCTGCTGTTATGAGAGATCAGCGACCTGAGGAGATGAC 561
Dy ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Dy 78 LeuArgGluArgLysGlnLeuValIleTyrGluGluLeuSerAspProGluGluAspAsp 97
QY 562 GAG 564
Dy ::::
Dy 98 Glu 98

RESULT 6
Q8WZ9 PRELIMINARY; PRT; 64 AA.
AC Q8WZ9;
DT 01-MAR-2002 (TREMELrel. 20, Created)
DT 01-MAR-2002 (TREMELrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMELrel. 20, Last annotation update)
DE BA564H18.1 (Synovial sarcoma, X breakpoint 2) (Fragment).
GN SSX2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RP SEQUENCE FROM N.A.
RA Whitehead S.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL596242; CAD18884.1; -.
FT NON TER 1
FT NON TER 64
SQ SEQUENCE 64 AA; 7014 MW; 43ADB72AF2FE9613 CRC64;

Alignment Scores:
Pred. No.: 7,93e-18 Length: 64
Score: 255.00 Matches: 48
Percent Similarity: 84.38% Conservative: 6
Best Local Similarity: 75.00% Mismatches: 10
Query Match: 24.88% Indels: 0
DB: 4 Gaps: 4

US-09-975-856-1 (1-576) x Q8WZ9 (1-64)
QY 283 GAACGTCTCAGATGACTTTTCGGCAGGCTCCAGAGAAATCTTCCGAGATCATGCCCAAG 342
Dy ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Dy 1 GluArgProGlnMetThrPheGlyArgLeuGlnGlyIleSerProLysIleMetProLys 20
QY 343 AAGCCACAGAGAGAAATGTTTGAAGAGAGTGCAGAGGCGATCTGGCCCAAAAT 402
Dy ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Dy 21 LysProAlaGluGluGlyAsnAspSerGluGluValProGluAlaSerGlyProGlnAsn 40
QY 403 CATGGGAAACAGCTGTGCCCGGGAAATCCCAATCTTGGAGAGATTAAACAGACA 462
Dy ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Dy 41 AspGlyLysGluLeuCysProGlyLysProThrThrSerGluLysIleHisGluArg 60
QY 463 TCTGACCCCAA 474
Dy ::::
Dy 61 SerGlyLysArg 64

RESULT 7
Q9NZK4 PRELIMINARY; PRT; 113 AA.
AC Q9NZK4;
DT 01-OCT-2000 (TREMELrel. 15, Created)
DT 01-OCT-2000 (TREMELrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMELrel. 15, Last annotation update)
DE SSX2 (Fragment).
GN SSX2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RP SEQUENCE FROM N.A.
RX MEDLINE=20210594; PubMed=10749136;
RA dos Santos N.R., Torenema R., de Vries T.J., Schreurs M.W.J.,
RA de Bruijn D.R.H., Kater-Baats E., Ruiter D.J., Adema G.J.,
RA van Muijen G.N.P., Geurts van Kessel A.;
RT "Heterogeneous expression of the SSX cancer/testis antigens in human
RT melanoma lesions and cell lines.";
RL Cancer Res. 60:1654-1662(2000).
DR EMBL; AF190791; AAF44724.1; -.
FT NON TER 1
SQ SEQUENCE 113 AA; 12399 MW; 981EBD852BA31DF8 CRC64;

Alignment Scores:
Pred. No.: 5,29e-17 Length: 113
Score: 247.50 Matches: 52
Percent Similarity: 50.88% Conservative: 6
Best Local Similarity: 45.61% Mismatches: 6
Query Match: 24.15% Indels: 50
DB: 4 Gaps: 1

US-09-975-856-1 (1-576) x Q9NZK4 (1-113)
QY 331 ATCATGCCACAGAGCCAGCAGAGAGAGAAATGTTTGAAGAGAGTCCAGAGCATCT 390
Dy ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Dy 1 IleMetProLysLysProAlaGluGlyAsnAspSerGluGluValProGluAlaSer 20
QY 391 GCGCCACAAATGATGGGAAACAGCTGTGCCCGGGAAATCCAGTACCTTGGAGAG 450
Dy ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Dy 21 GlyProGlnAsnAspGlyLysGluLeuCysProGlyLysProThrThrSerGluLys 40
QY 451 ATTAACAAGACATCTGGA-
Dy ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Dy 41 IleHisGluArgSerGlyAsnArgGluAlaGlnGluLysGluGluArgGlyThrAla 60
QY 468 -----
Dy 61 HisArgTyrSerSerGlnAsnThrHisAsnIleGlyArgPheSerLeuSerThrSerMet 80
QY 469 -----
Dy 81 GlyAlaValHisGlyThrProLysThrIleThrHisAsnArgAspPro-LysGlyGlyAs 100
```

```
Qy 484 CATGCTGGAGCCACAGACTGCGTGAGAGAAAGCAGCTGG 523
|||||
Db 100 nMetProGlyProThrAspCysValArgGluAenSerTrp 113
|||||

RESULT 8
Q80ZT4
ID Q80ZT4 PRELIMINARY; PRT; 128 AA.
AC Q80ZT4;
DT 01-JUN-2003 (TRENBLrel. 24, Created)
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Similar to RIKEN cDNA 4930414C09 gene.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testicle;
RA Strausberg R.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC048441; A48441.1;
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001909; KRAB.
DR InterPro; IPR003655; KRAB_related.
DR Pfam; PF01352; KRAB; 1.
DR SMART; SM00349; KRAB; 1.
DR PROSITE; PS08005; KRAB; 1.
DR PROSITE; PS08006; KRAB_RELATED; 1.
SQ SEQUENCE 128 AA; 15290 MW; F88614D1CBFF3B70 CRC64;

Alignment Scores:
Pred. No.: 1.5e-13 Length: 128
Score: 214.50 Matches: 56
Percent Similarity: 45.00% Conservative: 25
Best Local Similarity: 31.11% Mismatches: 36
Query Match: 20.93% Indels: 63
DB: 11 Gaps: 5

US-09-975-856-1 (1-576) x Q80ZT4 (1-128)
Qy 25 AGGAGACCCAGGATGATGCTCAAAATATCAGAGAAGTTACGAAAGGCTTCGATGATATT 84
::: ||| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
Db 9 LysAsnProMetGluValLeuTyrGluProLysAsnLeuCysLysAlaPheGlnAspIle 28
::: ||| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
Qy 85 GCCAAATCTCTCTTAAGAAAGAGTGGGAAAGATGAATCTCCGGAGAAAATCGTCTAT 144
::: ||| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
Db 29 SerThrTyrPheSerAspGluTyrGlyLysLeuThrGlnTyrGlnLysSerAlaTyr 48
::: ||| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
Qy 145 GTGTATATGAAGCTAAACTAGTCTACATAAAGTCTTCAAGGTCCACCTCCCA 204
::: ||| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
Db 49 ValTyrMetLysArgAsnTyrIleArgMetThrAspLeuGlyValThrValAsnGlnPro 68
::: ||| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
Qy 205 CCTTTCATGCGTAGTAAACCGGCTGCAGACTTCCACGGGAATGATTTGGTAACGATCGA 264
::: ||| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
Db 69 ValPheMetArg-----GlyLysGluGln 76
::: ||| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
Qy 265 AACACAGGAATCAGGTTGACGTCCTCAGATGACTTTCGACGCTCCAGAAATCTTC 324
::: ||| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
Db 77 AspLysGlnSerLeuValGlu----- 83
::: ||| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
Qy 325 CCGAAGATCATGCCCAAGAACCCAGACGAGAGAAATGTTTGAAGGAAGTCCAGAG 384
::: ||| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
Db 84 -----GlyIle-----GluValHisAsp 89
::: ||| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
Qy 385 GCATCTGCCCCACAAAATGATGGGAAACAGCTGTGCCCCCGGGAAATCCAAAGTACCTTG 444
::: ||| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
Db 90 Ser----- 90
::: ||| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
Qy 445 GAGAAGATTAAACAGACATCTGGACCCCAAGGGGAAACATGCTGCGACCCACAGACTG 504
::: ||| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
```

```

Db 121 -----Glytle---GluValHisAsp 126
QY 385 GCATCTGCCCAAAATGATGGAAACAGCTGTGCCCCCGGAAATCAAGTACCTTC 444
Db 127 Ser:::
QY 445 GAGAAGATTAAACAGACATCTGGACCCCAAAAGGGGAAACATGCTGGACCCACAGACTG 504
Db 128 -----AspGluThrSerGlylearg-----ValAsnValTrpSerHisArgGlu 142
QY 505 CGTCAGAGAAAGCAGCTGGTGGTTTATCAAGAGATCAGCCAGCTGGAGGAGATGACGAG 564
Db 143 ArgGluArgLysTyrArgValIleTyrGluGluLysSerAspProGluGluGluAsp 162

RESULT 10
Q9CPU1 PRELIMINARY; PRT; 170 AA.
AC Q9CPU1;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE 4930414C09RIK protein.
GN 4930414C09RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Musinae; Mus.
OX NCBI_TaxID=10090;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=21086660; PubMed=1217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleisichmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikolaio I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staebli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Montbaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohseki S.,
RA Hayashizaki Y.
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 403:685-690(2001).
DR EMBL; AK015135; BAB29722.1; -.
DR EMBL; AK006218; BAB24465.1; -.
DR MGP; MGI:1915235; 4930414C09RIK.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0003676; P:nucleic acid binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001909; KRAB.
DR InterPro; IPR003855; KRAB-related.
DR Pfam; PF01352; KRAB; 1.
DR SMART; SM00349; KRAB; 1.
DR PROSITE; PS50805; KRAB; 1.
DR PROSITE; PS50806; KRAB-RELATED; 1.
SQ SEQUENCE 170 AA; 19636 MW; F63F702BFDF7227 CRC64;

Alignment Scores:
Pred. No.: 3.24e-13 Length: 170
Score: 211.50 Matches: 58
Percent Similarity: 47.34% Conservative: 22
Best Local Similarity: 34.32% Mismatches: 62
Query Match: 20.63% Indels: 27
DB: 11 Gaps: 5

```

```

US-09-975-856-1 (1-576) x Q9CPU1 (1-170)
QY 67 AAGGCTTCGATGATATTCGCAATATCTCTTAAGAAAGAGTGGGAAAGATGAATCC 126
Db 23 LysAlaPheGlnAspIleSerThrTyrPheSerAspGluGluTrpGlyLysLeuThrGln 42
QY 127 TCGGAGAAAATCGTCTATGTATGAAGCTAAAGTAACTATGAGGTCACTGACTAAACTAGGT 186
Db 43 TrpGlnLysSerAlaTyrValTyrMetLysArgAsnTyrIleArgMetThrAspLeuGly 62
QY 187 TTCAAGGTCAACCTCCACCTTTCATGCGTAGTAGTAAAGGGGTGCAGAC-----TTC 237
Db 63 ValThrValAsnGlnProValPheMetArgGlyLysGluGlnAlaLysGlnSerLeuVal 82
QY 238 CACGGGAATGATTTGGTAAACGATCGAACCACAGGATCAGGTGAGCGTCTCCAGATG 297
Db 83 GluGlyIleGluVal-----HisAspSerGluAspGluCysPheGluGly 97
QY 298 ACTTTCGGCAGCCTCCAGAGAATCTTCCGAGATCATGCCCCAAGAACCCAGCAGAGAA 357
Db 98 SerPheGly-----ValThrProIleLysArgMetLysLeu 109
QY 358 GAAATGGTTTGAAGGAAGTCCAGAGCGCATCTGCCCCACAAAATGATGGGAAACAGCTG 417
Db 110 ThrSerValThrIleSerPheHisAsnValGluGlySerLeuAlaSerGlyGluAsnAsp 129
QY 418 TGCCCCCGGGGAAATCCAACTACCTTGGAGAAGATTAACAAGACATCTGGACCCCAAAAGG 477
Db 130 Cys-----AsnLeuAlaGluThrGlyGlyIleGln--- 139
QY 478 GGGAAACATGCTGCGACCCACAGACTGCTGAGAGAAGAGCAGCTGTTTATGAAGAG 537
Db 140 ---ValAsnValTrpSerHisArgLeuArgGluArgTyrArgValIleTyrSerGlu 158
QY 538 ATCAGCGACCCCTGAGGAAGATGACGAG 564
Db 159 IleSerAspThrGluGluGluAsp 167

RESULT 11
O75101 PRELIMINARY; PRT; 117 AA.
ID O75101;
AC O75101;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE SSX-HSTT (Fragment).
GN SSX-HSTT.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN 11
RP SEQUENCE FROM N.A.
RA Sonobe H., Takeuchi T.;
RT "SSX-HSTT.";
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB012575; BAA32799.1; -.
FT NON_TER 1
FT NON_TER 117
SQ SEQUENCE 117 AA; 12611 MW; C9F9ADFCD0AD4542 CRC64;

Alignment Scores:
Pred. No.: 6.39e-08 Length: 117
Score: 160.50 Matches: 34
Percent Similarity: 64.62% Conservative: 8
Best Local Similarity: 52.31% Mismatches: 16
Query Match: 15.68% Indels: 7
DB: 4 Gaps: 2
US-09-975-856-1 (1-576) x O75101 (1-117)
QY 241 GGAATGATTTGGTAAACGATCGAACCACAGGATCAGGTGTAACGCTCT-----CAG 294

```


OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Wei Y., Sun M., Wang J., Zhu X., Larsson C., Dwight T., Xie Y.,
RA Nilsson G., Larsson O.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY138491; AAN39533.1; -.
FT NON_TER 1
FT NON_TER 39
SQ SEQUENCE 39 AA; 4336 MW; 8FC179F6C8C7E0D CRC64;

Alignment Scores:
Pred. No.: 1.51e-05 Length: 39
Score: 137.00 Matches: 25
Percent Similarity: 87.68% Conservative: 4
Best Local Similarity: 75.76% Mismatches: 4
Query Match: 13.37% Indels: 0
DB: 4 Gaps: 0

US-09-975-856-1 (1-576) x Q81ZG7 (1-39)

QY 322 TTCCCGAAGATCATGCCCAAGACGACGAGGAGAAATGTTTGAAGGAAGTGCCA 381
::: :
Db 7 TyrAspGlnIleMetProLysLysProAlaGluGluGlyAsnAspSerGluGluValPro 26

QY 382 GAGGATCTGCCCCACAAAATGATGGGAAACAGCTGTGC 420
|||||
Db 27 GluAlaSerGlyProGlnAsnAspGlyLysGluLeuCys 39

Search completed: March 31, 2004, 13:59:50
Job time : 50 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 31, 2004, 15:16:19 ; Search time 2745 Seconds

(without alignments)
6266.162 Million cell updates/sec

Title: US-09-975-856-1

Perfect score: 576

Sequence: 1 ATGAACGGAGACGACGCTT.....ATGACGAGTAACCTCCCTCG 576

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST.*

1: em_estba.*

2: em_esthum.*

3: em_estin.*

4: em_estmu.*

5: em_estov.*

6: em_estpl.*

7: em_estro.*

8: em_hic.*

9: gb_estl.*

10: gb_est2.*

11: gb_hic.*

12: gb_est3.*

13: gb_est4.*

14: gb_est5.*

15: em_estfun.*

16: em_estom.*

17: em_gss_hum.*

18: em_gss_inv.*

19: em_gss_pln.*

20: em_gss_vrt.*

21: em_gss_fun.*

22: em_gss_mam.*

23: em_gss_mus.*

24: em_gss_pro.*

25: em_gss_rod.*

26: em_gss_phg.*

27: em_gss_vrl.*

28: gb_gss1.*

29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	496.6	86.2	979	12	BM806411
2	495	85.9	836	13	BU194624
3	495	85.9	887	13	BU161779
4	491	85.2	863	13	BQ222907

5	488.6	84.8	1001	12	BM471266
6	481.6	83.6	781	13	BQ434972
7	481.6	83.6	873	13	BU170242
8	481.6	83.6	888	13	BQ432374
9	479.2	83.2	1135	10	BE535379
10	471.6	81.9	602	14	BE535379
11	470.8	81.7	950	13	BQ231752
12	454.6	78.9	793	12	BE560469
13	447	77.6	922	13	BQ229996
14	423.8	73.6	573	10	BE408883
15	423.4	73.5	867	10	BE891434
16	417.6	72.5	627	14	CD767542
17	408.8	71.0	589	10	BE410950
18	402	69.8	943	13	BQ229064
19	400.8	69.6	948	12	BQ333981
20	394.4	68.5	630	10	BE408978
21	394.4	68.5	660	14	CB156144
22	394.4	68.5	759	10	BE384545
23	394.4	68.5	798	10	BE410794
24	388.6	67.5	866	12	EG478295
25	383.4	66.6	636	10	BE390290
26	383	66.5	678	10	BF211314
27	382.4	66.4	757	12	BQ104299
28	377.6	65.6	1053	13	BQ241117
29	373	64.8	848	10	BE387586
30	371.4	64.5	961	13	BU189304
31	370.8	64.4	662	10	BE391023
32	367.6	63.8	1057	12	EG476029
33	360.2	62.5	655	12	BI811262
34	360	62.5	988	10	BE729870
35	360	62.5	1007	12	BQ396977
36	357.6	62.1	1006	13	BU189362
37	340.6	59.1	533	10	BF184274
38	339.4	58.9	958	13	BQ232177
39	330.6	57.4	623	10	BE896093
40	326	56.6	367	10	BE866727
41	324	56.2	927	13	BU178938
42	322	55.9	486	9	AA312651
43	322	55.9	896	10	BF212335
44	319.8	55.5	484	14	CB143725
45	317	55.0	820	10	BE777476

ALIGNMENTS

RESULT 1

BM806411

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BM806411 979 bp mRNA linear EST 05-MAR-2002
AGENCOURT_6542819 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5548810
5', mRNA sequence.

BM806411

BM806411.1 GI:19123234

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

NIH-MGC <http://mgi.nci.nih.gov/>

1 (bases 1 to 979)

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: ATCC/DCTD/DTF

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLAM12258 row: f column: 11

High quality sequence stop: 622.

FEATURES source
1. .979
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5548810"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_72"
/note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2 kb. Library constructed by Life
Technologies."

NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: ATCC/DCTD/DTF
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
DNA Distribution: Agencourt Bioscience Corporation
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: L1AM13525 row: f column: 16
High quality sequence stop: 624.
Location/Qualifiers
1. .836
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6165903"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_72"
/note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2 kb. Library constructed by Life
Technologies."

ORIGIN

Query Match 86.2%; Score 496.6; DB 12; Length 979;
Best Local Similarity 91.5%; Pred. No. 1.8e-131;
Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
QY 1 ATGAACGAGACGAGCCTTTGCAAGGAGACCCAGGAGATGATCTCAAAATATCAGAGAAG 60
Db 65 ATGAACGAGACGAGCCTTTGCAAGGAGACCCAGGAGATGATCTCAAAATATCAGAGAAG 124
QY 61 TTACGAAAGCCTTCGATGATATTTGCCAAATACCTTCTTAAGAAAGAGTGGGAAAAGATG 120
Db 125 ATCCAAAAGCCTTCGATGATATTTGCCAAATACCTTCTTAAGAAAGAGTGGGAAAAGATG 184
QY 121 AAATCCTCGGAGAAATCGTCTATGTATATGAGCTAACTATGAGTCACTAA 180
Db 185 AAAGCCTCGGAGAAATCTTCTATGTATATGAGGAAAGATGAGGCTATGACTAA 244
QY 181 CTAGGTTTCAAGGTCACCCCTCCACACCTTTCATGCTAGTAAACGGGCTGACAGACTTCCAC 240
Db 245 CTAGGTTTCAAGGTCACCCCTCCACACCTTTCATGCTAGTAAACGGGCTGACAGACTTCCAC 304
QY 241 GGGATGATTTGGTACGATCGAACCACAGAAATCAGGTTGAAAGTCTCTCAGTACT 300
Db 305 GGGATGATTTGGTACGATCGAACCACAGAAATCAGGTTGAAAGTCTCTCAGTACT 364
QY 301 TTCCGACGCTCCAGAGAAATCTTCCGAGATCATGCCCAAGAGCCAGCAGAGGAGAA 360
Db 365 TTCCGACGCTCCAGAGAAATCTTCCGAGATCATGCCCAAGAGCCAGCAGAGGAGAA 424
QY 361 AATGGTTGAAGGAGTGCAGAGGATCTGGGCCACAAATGATGGGAAACAGCTGTGC 420
Db 425 AATGATTCGGAGGAAGTGCAGAGGATCTGGGCCACAAATGATGGGAAACAGCTGTGC 484
QY 421 CCCCAGGAAATCCAACTACCTCTGAGAGATTAACAGACATCTGGACCCAAAGGGGG 480
Db 485 CCCCAGGAAATCCAACTACCTCTGAGAGATTAACAGAGATCTGGACCCAAAGGGGG 544
QY 481 AAACATGCTGGACCCACAGATCGGTGAGAGAAAGCAGCTGGTGTATGAGAGATC 540
Db 545 GAAATGCTGGACCCACAGATCGGTGAGAGAAACAGCTGGTGTATGAGAGATC 604
QY 541 AGCCACCTGAGGAGATGACGAGTAATCTCCCTC 575
Db 605 AGCCACCTGAGGAGATGACGAGTAATCTCCCTC 639

RESULT 2
BUI94624
LOCUS
DEFINITION AGENCOURT_7969283 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6165903
5', mRNA sequence.
ACCESSION BUI94624
VERSION BUI94624.1 GI:22708608
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (Bases 1 to 836)

FEATURES source

Query Match 85.9%; Score 495; DB 13; Length 836;
Best Local Similarity 91.3%; Pred. No. 4.9e-131;
Matches 525; Conservative 0; Mismatches 50; Indels 0; Gaps 0;
QY 1 ATGAACGAGACGAGCCTTTGCAAGGAGACCCAGGAGATGATCTCAAAATATCAGAGAAG 60
Db 74 ATGAACGAGACGAGCCTTTGCAAGGAGACCCAGGAGATGATCTCAAAATATCAGAGAAG 133
QY 61 TTACGAAAGCCTTCGATGATATTTGCCAAATACCTTCTTAAGAAAGAGTGGGAAAAGATG 120
Db 134 ATCCAAAAGCCTTCGATGATATTTGCCAAATACCTTCTTAAGAAAGAGTGGGAAAAGATG 193
QY 121 AAATCCTCGGAGAAATCGTCTATGTATATGAGCTAACTATGAGTCACTAA 180
Db 194 AAAGCCTCGAGAAATCTTCTATGTATATGAGGAAAGATGAGGCTATGACTAA 253
QY 181 CTAGGTTTCAAGGTCACCCCTCCACACCTTTCATGCTAGTAAACGGGCTGACAGACTTCCAC 240
Db 254 CTAGGTTTCAAGGTCACCCCTCCACACCTTTCATGCTAGTAAACGGGCTGACAGACTTCCAC 313
QY 241 GGGATGATTTGGTACGATCGAACCACAGAAATCAGGTTGAAAGTCTCTCAGTACT 300
Db 314 GGGATGATTTGGTACGATCGAACCACAGAAATCAGGTTGAAAGTCTCTCAGTACT 373
QY 301 TTCCGACGCTCCAGAGAAATCTTCCGAGATCATGCCCAAGAGCCAGCAGAGGAGAA 360
Db 374 TTCCGACGCTCCAGAGAAATCTTCCGAGATCATGCCCAAGAGCCAGCAGAGGAGAA 433
QY 361 AATGGTTGAAGGAGTGCAGAGGATCTGGGCCACAAATGATGGGAAACAGCTGTGC 420
Db 434 AATGATTCGGAGGAAGTGCAGAGGATCTGGGCCACAAATGATGGGAAACAGCTGTGC 493
QY 421 CCCCAGGAAATCCAACTACCTCTGAGAGATTAACAGACATCTGGACCCAAAGGGGG 480
Db 494 CCCCAGGAAATCCAACTACCTCTGAGAGATTAACAGAGATCTGGACCCAAAGGGGG 553
QY 481 AAACATGCTGGACCCACAGATCGGTGAGAGAAAGCAGCTGGTGTATGAGAGATC 540
Db 554 GAAATGCTGGACCCACAGATCGGTGAGAGAAACAGCTGGTGTATGAGAGATC 613
QY 541 AGCCACCTGAGGAGATGACGAGTAATCTCCCTC 575
Db 614 AGCCACCTGAGGAGATGACGAGTAATCTCCCTC 648

```
RESULT 3
BUI61779      887 bp  mRNA  linear  EST 04-SEP-2002
LOCUS         5', mRNA sequence.
DEFINITION    BUI61779 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6168477
ACCESSION     BUI61779
VERSION       BUI61779.1 GI:22675689
KEYWORDS      EST.
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens
REFERENCE     1 (bases 1 to 887)
AUTHORS      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE        Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
JOURNAL      NIH-MGC http://mgc.nci.nih.gov/.
COMMENT      National Institutes of Health, Mammalian Gene Collection (MGC)
              Unpublished (1999)
              Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-remail.nih.gov
              Tissue Procurement: ATCC/DCTD/DTF
              cDNA Library Preparation: Life Technologies, Inc.
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
              DNA Sequencing by: Agencourt Bioscience Corporation
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              Plate: LUAM13532 row: a column: 22
              High quality sequence stop: 739.
              Location/Qualifiers
                1..887
                  /organism="Homo sapiens"
                  /mol_type="mRNA"
                  /db_xref="taxon:9606"
                  /clone="IMAGE:6168477"
                  /tissue_type="melanotic melanoma"
                  /lab_host="DH10B (phage-resistant)"
                  /clone_lib="NIH MGC 72"
                  /note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
                  Site 2: SalI; Cloned unidirectionally. Primer: Oligo dr.
                  Average insert size 2 kb. Library constructed by Life
                  Technologies."
FEATURES
  source
    Query Match      85.9%; Score 495; DB 13; Length 887;
    Best Local Similarity 91.3%; Pred. No. 5, 1e-13;
    Matches 525; Conservative 0; Mismatches 50; Indels 0; Gaps 0;
  QY 1 ATGAACGGAGACGACGCGCTTTGCAAGGAGACCCAGGATGCTCAAAATATCAGAGAAG 60
  DB 75 ATGAACGGAGACGACGCGCTTTGCAAGGAGACCCAGGATGCTCAAAATATCAGAGAAG 134
  QY 61 TTACGAAAGGCGCTTCGATGATATTCGCAAAATATCTCTTAAGAAGAGTGGGAAAAGATG 120
  DB 135 ATCCAAAGGCGCTTCGATGATATTCGCAAAATATCTCTTAAGAAGAGTGGGAAAAGATG 194
  QY 121 AAATCCTCGGAGAAATCGTCTATGTGTATATGAAGCTAAACTATGAGGTATGACTATAA 180
  DB 195 AAAGCCTCAGAGAAATCTTCTATGTGTATATGAAGAGAAAGTATGAGGCTATGACTATAA 254
  QY 181 CTAGGTTTCAAGGTCACCGCTCCACCTTTCATGCGTACTTAACCGGCTGCAGCTCCAC 240
  DB 255 CTAGGTTTCAAGGTCACCGCTCCACCTTTCATGCGTACTTAACCGGCGCCAGAGCTCCAG 314
  QY 241 GCGAATGATTTTGGTAAACGATCGAAACACACAGGAATCAGGTTGAACGTCCTCAGATGACT 300
  DB 315 GCGAATGATTTTGGTAAACGATCGAAACACACAGGAATCAGGTTGAACGTCCTCAGATGACT 374
  QY 301 TTCGGCAGCCTCCAGAGATCTTCCGGAAGATCATGCCAGAGACCCAGCAGAGAGAA 360
  DB 375 TTCGGCAGCCTCCAGAGATCTTCCGGAAGATCATGCCAGAGACCCAGCAGAGAGAA 434
  QY 361 AATGGTTTGAAGGAAGTGCAGAGGCATCTGCGCCACAAAAATGATGGGAAAACAGCTGTGC 420
```

```
DB 435 AATGATTCGGAGGAAGTCCAGAAAGCATCTGGCCCCAACAAATGATGGAAAGAGCTGTGC 494
QY 421 CCCCGGGAAATCAAAGTACCTTCGAGAGAAATTAAACAGACATCTGGACCCCAAAAGGGGG 480
DB 495 CCCCGGGAAATCAAAGTACCTTCGAGAGAAATTAAACAGAGATCTGGACCCCAAAAGGGGG 554
QY 481 AAACATGCTGGACCCAGACAGCTCGGTGAGAGAAAGCAGCTGGTGTATGAAGAGATC 540
DB 555 GAACATGCTGGACCCAGACAGCTCGGTGAGAGAAAGCAGCTGGTGTATGAAGAGATC 614
QY 541 AGCGACCTCGAGGAAGATGACGAGTAACCTCCCTC 575
DB 615 AGCGACCTCGAGGAAGATGACGAGTAACCTCCCTC 649

BQ222907      863 bp  mRNA  linear  EST 02-MAY-2002
LOCUS         5', mRNA sequence.
DEFINITION    BQ222907 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6095680
ACCESSION     BQ222907
VERSION       BQ222907.1 GI:20404307
KEYWORDS      EST.
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens
REFERENCE     1 (bases 1 to 863)
AUTHORS      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE        Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
JOURNAL      NIH-MGC http://mgc.nci.nih.gov/.
COMMENT      National Institutes of Health, Mammalian Gene Collection (MGC)
              Unpublished (1999)
              Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-remail.nih.gov
              Tissue Procurement: ATCC/DCTD/DTF
              cDNA Library Preparation: Life Technologies, Inc.
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
              DNA Sequencing by: Agencourt Bioscience Corporation
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              Plate: LUAM13368 row: h column: 17
              High quality sequence stop: 625.
              Location/Qualifiers
                1..863
                  /organism="Homo sapiens"
                  /mol_type="mRNA"
                  /db_xref="taxon:9606"
                  /clone="IMAGE:6095680"
                  /tissue_type="melanotic melanoma"
                  /lab_host="DH10B (phage-resistant)"
                  /clone_lib="NIH MGC 72"
                  /note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
                  Site 2: SalI; Cloned unidirectionally. Primer: Oligo dr.
                  Average insert size 2 kb. Library constructed by Life
                  Technologies."
FEATURES
  source
    Query Match      85.2%; Score 491; DB 13; Length 863;
    Best Local Similarity 91.2%; Pred. No. 7e-130;
    Matches 521; Conservative 0; Mismatches 50; Indels 0; Gaps 0;
  QY 5 ACCGAGACGACGCGCTTTCGAGAGAGACCCAGGATGATGCTCAAAATATCAGAGAAGTAC 64
  DB 1 ACCGAGACGACGCGCTTTCGAGAGAGACCCAGGATGATGCTCAAAATATCAGAGAAGTAC 60
  QY 65 GAAAGGCGCTTCGATGATTTGCAAAATATCTTCTTAAGAAAGAGTGGGAAAAGATGAAT 124
  DB 61 AAAAGGCGCTTCGATGATTTGCAAAATATCTTCTTAAGAAAGAGTGGGAAAAGATGAAG 120
  QY 125 CCTCGGAGAAATCGTCTATGTGTATATGAAGCTAACTATGAGGTATGACTTAACTAG 184
  DB 121 CCTCAGAGAAATCTTCTATGTGTATATGAAGAGAAAGTATGAGGCTATGACTTAACTAG 180
```

```

QY 185 GTTCAAGGTCAACCTCCACCTTTTCATGCTAGTAAACGGGCTGCAGACTTCCACGGGA 244
DB 181 GTTCAAGGCCACCTCCACCTTTTCATGCTAGTAAACGGGCGGAGACTTCCAGGGGA 240
QY 245 ATGATTTGGTAACGATCGAACCAACAGAAATCAGTTGAAAGTCTCAGATGACTTTTCG 304
DB 241 ATGATTTGGTAATGACCTTAACCGTGGGAATCAGTTGAAAGTCTCAGATGACTTTTCG 300
QY 305 GCAGGCTCCAGAGATCTTCCGGAAGATCATGCCCAAGAACCCAGCAGAGGAAGAAATG 364
DB 301 GCAGGCTCCAGGATCTTCCGGAAGATCATGCCCAAGAACCCAGCAGAGGAAGAAATG 360
QY 365 GTTGAAGAAAGTCCAGAGGATCTGCCCAACAAATGATGGAAACAGCTGTGCCCCC 424
DB 361 ATTCGAGGAAGTGCAGAGCATCTGCCCAACAAATGATGGAAACAGCTGTGCCCCC 420
QY 425 CGGGAATCCCAAGTACCTTGGAGAGATTAACAGAGCATCTGCCCAACAAAGGGGAAAC 484
DB 421 CGGGAATCCCAAGTACCTTGGAGAGATTAACAGAGCATCTGCCCAACAAAGGGGAAAC 480
QY 485 ATGCTCGACACCAAGCTGCTGTGAGAGAAAGCAGCTGTGTTTATGAAGAGATCAGCG 544
DB 481 ATGCTCGACACCAAGCTGCTGTGAGAGAAAGCAGCTGTGTTTATGAAGAGATCAGCG 540
QY 545 ACCCTGAGAGATCAAGAGTAACCTCCCTC 575
DB 541 ACCCTGAGAGATCAAGAGTAACCTCCCTC 571

```

RESULT 5

BM471266

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BM471266 1001 bp mRNA linear EST 05-FEB-2002
 AGENCOURT_6478381 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5563101
 5', mRNA sequence.

BM471266

BM471266.1 GI:18520308

EST.

Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1001)

NIH-MGC http://mgc.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabbs@mail.nih.gov

Tissue Procurement: ATCC/DCTD/DTF

cDNA library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LNL at:

http://image.lnl.gov

Plate: L1AM1293 row: i column: 22

High quality sequence start: 258

High quality sequence stop: 436.

Location/Qualifiers

1. .1001

/organism="Homo sapiens"

/db_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5563101"

/issue_type="melanotic melanoma"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH MGC 72"

/notes="Organ: skin; Vector: pCMV-SPORT6; Site 1: Not I;

Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.

Average insert size 2 kb. Library constructed by Life

Technologies."

FEATURES

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

Best Local Similarity 90.6%; Pred. No. 3.7e-129;
 Matches 521; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

```

QY 1 ATGAACGGAGACGACCCCTTTGCAAGGAGACCCAGGATGATGCTCAATATACAGAGAG 60
DB 78 ATGAACGGAGACGACCCCTTTGCAAGGAGACCCAGGATGATGCTCAATATACAGAGAG 137
QY 61 TTACCAAAAGCCCTTCGATGATATTCGCAAAATCTTCTCTAAGAAAGAGTGGGAAAGATG 120
DB 138 ATCCAAAAGCCCTTCGATGATATTCGCAAAATCTTCTCTAAGAAAGAGTGGGAAAGATG 197
QY 121 AATCTCTCGAGAAATCTGCTATGTATATGAAGCTAAACTATGAGTCACTGACTTAA 180
DB 198 AGAGCCTCAGAGAAATCTTCTATGTATATGAAGAGAAATGATGAGGCTATGACTAAA 257
QY 181 CTAGGTTTCAAGGTCAACCTCCACCTTTTCATGCTAGTAAACGGGCTGCAGACTTCCAC 240
DB 258 CTAGGTTTCAAGGCCACCTCCACCTTTTCATGCTAGTAAACGGGCTGCAGACTTCCAG 317
QY 241 GGAATGATTTGTTGTAACGATCGAAACACAGAGATCAGTTGAGCTCTCAGATGACT 300
DB 318 GGAATGATTTGTTGTAACGATCGAAACACAGAGATCAGTTGAGCTCTCAGATGACT 377
QY 301 TTCGGCAGCTCCAGAGAAATCTTCCGAAAGATCATGCCCAAGAACCCAGCAGAGGAAGA 360
DB 378 TTCGGCAGATCCAGGGAATCTCCCGAAGATCATGCCCAAGAACCCAGCAGAGGAAGA 437
QY 361 AATGTTTGAAGGAAGTCCAGAGGATCTGCGCCCAAAAATGATGGGAAACAGCTGTGC 420
DB 438 AATGTTTGAAGGAAGTCCAGAGGATCTGCGCCCAAAAATGATGGGAAAGAGCTGTGC 497
QY 421 CCCCCGGGAAATCCCAAGTACCTTGGAGAGAGATTAAACAGACATCTGAGCCCAAAAGGGG 480
DB 498 CCCCCGGGAAACCAACTACTCTCTGAGAGATTCAGAGAGATCTGAGCCCAAAAGGGG 557
QY 481 AAACATGCTCGACCCACAGACTGCTGTGAGAGAAAGCAGCTGTGTTTATGAAGAGATC 540
DB 558 GAACATGCTCGACCCACAGACTGCTGTGAGAGAAAGCAGCTGTGTTTATGAAGAGATC 617
QY 541 AGGACCTCTGAGGAAGATGACGAGTAACCTCCCTC 575
DB 618 AGGACCTCTGAGGAAGATGACGAGTAACCTCCCTC 652

```

RESULT 6

BM434972

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BM434972 781 bp mRNA linear EST 24-MAY-2002
 AGENCOURT_7896851 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6159943
 5', mRNA sequence.

BM434972

BM434972.1 GI:21174048

EST.

Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 781)

NIH-MGC http://mgc.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabbs@mail.nih.gov

Tissue Procurement: ATCC/DCTD/DTF

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LNL at:

http://image.lnl.gov

Plate: L1AM13509 row: n column: 08

High quality sequence stop: 634.

Location/Qualifiers

1. .781

source

source

source

source

source

source

QY 185 GTTCAAGGTCAACCTCCACCTTTTCATGCTAGTAAACGGGCTGCAGACTTCCACGGGA 244

DB 181 GTTCAAGGCCACCTCCACCTTTTCATGCTAGTAAACGGGCGGAGACTTCCAGGGGA 240

QY 245 ATGATTTGGTAACGATCGAACCAACAGAAATCAGTTGAAAGTCTCAGATGACTTTTCG 304

DB 241 ATGATTTGGTAATGACCTTAACCGTGGGAATCAGTTGAAAGTCTCAGATGACTTTTCG 300

QY 305 GCAGGCTCCAGAGATCTTCCGGAAGATCATGCCCAAGAACCCAGCAGAGGAAGAAATG 364

DB 301 GCAGGCTCCAGGATCTTCCGGAAGATCATGCCCAAGAACCCAGCAGAGGAAGAAATG 360

QY 365 GTTGAAGAAAGTCCAGAGGATCTGCCCAACAAATGATGGAAACAGCTGTGCCCCC 424

DB 361 ATTCGAGGAAGTGCAGAGCATCTGCCCAACAAATGATGGAAACAGCTGTGCCCCC 420

QY 425 CGGGAATCCCAAGTACCTTGGAGAGATTAACAGAGCATCTGCCCAACAAAGGGGAAAC 484

DB 421 CGGGAATCCCAAGTACCTTGGAGAGATTAACAGAGCATCTGCCCAACAAAGGGGAAAC 480

QY 485 ATGCTCGACACCAAGCTGCTGTGAGAGAAAGCAGCTGTGTTTATGAAGAGATCAGCG 544

DB 481 ATGCTCGACACCAAGCTGCTGTGAGAGAAAGCAGCTGTGTTTATGAAGAGATCAGCG 540

QY 545 ACCCTGAGAGATCAAGAGTAACCTCCCTC 575

DB 541 ACCCTGAGAGATCAAGAGTAACCTCCCTC 571

QY 545 ACCCTGAGAGATCAAGAGTAACCTCCCTC 575

DB 541 ACCCTGAGAGATCAAGAGTAACCTCCCTC 571

QY 545 ACCCTGAGAGATCAAGAGTAACCTCCCTC 575

DB 541 ACCCTGAGAGATCAAGAGTAACCTCCCTC 571

QY 545 ACCCTGAGAGATCAAGAGTAACCTCCCTC 575

DB 541 ACCCTGAGAGATCAAGAGTAACCTCCCTC 571

QY 545 ACCCTGAGAGATCAAGAGTAACCTCCCTC 575

DB 541 ACCCTGAGAGATCAAGAGTAACCTCCCTC 571

QY 545 ACCCTGAGAGATCAAGAGTAACCTCCCTC 575

DB 541 ACCCTGAGAGATCAAGAGTAACCTCCCTC 571

QY 545 ACCCTGAGAGATCAAGAGTAACCTCCCTC 575

DB 541 ACCCTGAGAGATCAAGAGTAACCTCCCTC 571

QY 545 ACCCTGAGAGATCAAGAGTAACCTCCCTC 575

DB 541 ACCCTGAGAGATCAAGAGTAACCTCCCTC 571

QY 545 ACCCTGAGAGATCAAGAGTAACCTCCCTC 575

DB 541 ACCCTGAGAGATCAAGAGTAACCTCCCTC 571

QY 545 ACCCTGAGAGATCAAGAGTAACCTCCCTC 575

DB 541 ACCCTGAGAGATCAAGAGTAACCTCCCTC 571

QY 545 ACCCTGAGAGATCAAGAGTAACCTCCCTC 575

DB 541 ACCCTGAGAGATCAAGAGTAACCTCCCTC 571

QY 545 ACCCTGAGAGATCAAGAGTAACCTCCCTC 575

DB 541 ACCCTGAGAGATCAAGAGTAACCTCCCTC 571

QY 545 ACCCTGAGAGATCAAGAGTAACCTCCCTC 575

DB 541 ACCCTGAGAGATCAAGAGTAACCTCCCTC 571

QY 545 ACCCTGAGAGATCAAGAGTAACCTCCCTC 575

DB 541 ACCCTGAGAGATCAAGAGTAACCTCCCTC 571

QY 545 ACCCTGAGAGATCAAGAGTAACCTCCCTC 575

DB 541 ACCCTGAGAGATCAAGAGTAACCTCCCTC 571

QY 545 ACCCTGAGAGATCAAGAGTAACCTCCCTC 575

DB 541 ACCCTGAGAGATCAAGAGTAACCTCCCTC 571

QY 545 ACCCTGAGAGATCAAGAGTAACCTCCCTC 575

DB 541 ACCCTGAGAGATCAAGAGTAACCTCCCTC 571

QY 545 ACCCTGAGAGATCAAGAGTAACCTCCCTC 575

DB 541 ACCCTGAGAGATCAAGAGTAACCTCCCTC 571

QY 545 ACCCTGAGAGATCAAGAGTAACCTCCCTC 575

DB 541 ACCCTGAGAGATCAAGAGTAACCTCCCTC 571

QY 545 ACCCTGAGAGATCAAGAGTAACCTCCCTC 575

DB 541 ACCCTGAGAGATCAAGAGTAACCTCCCTC 571

QY 545 ACCCTGAGAGATCAAGAGTAACCTCCCTC 575

DB 541 ACCCTGAGAGATCAAGAGTAACCTCCCTC 571

QY 545 ACCCTGAGAGATCAAGAGTAACCTCCCTC 575

DB 541 ACCCTGAGAGATCAAGAGTAACCTCCCTC 571

QY 545 ACCCTGAGAGATCAAGAGTAACCTCCCTC 575

DB 541 ACCCTGAGAGATCAAGAGTAACCTCCCTC 571

QY 545 ACCCTGAGAGATCAAGAGTAACCTCCCTC 575

DB 541 ACCCTGAGAGATCAAGAGTAACCTCCCTC 571

QY 545 ACCCTGAGAGATCAAGAGTAACCTCCCTC 575

DB 541 ACCCTGAGAGATCAAGAGTAACCTCCCTC 571

QY 545 ACCCTGAGAGATCAAGAGTAACCTCCCTC 575

DB 541 ACCCTGAGAGATCAAGAGTAACCTCCCTC 571

QY 545 ACCCTGAGAGATCAAGAGTAACCTCCCTC 575

DB 541 ACCCTGAGAGATCAAGAGTAACCTCCCTC 571

QY 545 ACCCTGAGAGATCAAGAGTAACCTCCCTC 575

DB 541 ACCCTGAGAGATCAAGAGTAACCTCCCTC 571

QY 545 ACCCTGAGAGATCAAGAGTAACCTCCCTC 575

DB 541 ACCCTGAGAGATCAAGAGTAACCTCCCTC 571

QY 545 ACCCTGAGAGATCAAGAGTAACCTCCCTC 575

DB 541 ACCCTGAGAGATCAAGAGTAACCTCCCTC 571

QY 545 ACCCTGAGAGATCAAGAGTAACCTCCCTC 575

DB 541 ACCCTGAGAGATCAAGAGTAACCTCCCTC 571

QY 545 ACCCTGAGAGATCAAGAGTAACCTCCCTC 575

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6159943"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC_72"
/note="Organ: skin; Vector: pCMV-SPORT6; Site:1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2 kb. Library constructed by Life
Technologies."

ORIGIN

Query Match 83.6%; Score 481.6; DB 13; Length 781;
Best Local Similarity 89.8%; Pred. No. 3.4e-127;
Matches 517; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 1 ATGAACGGAGACGACCCCTTTCAGAGGAGACCCAGGATGATGCTCAATATCAGAGAAG 60
Db ATGAACGGAGACGACCCCTTTCAGAGGAGACCCAGGATGATGCTCAATATCAGAGAAG 136
QY 61 TTACGAAAGCCCTTCGATGATATTCGCAATATCTCTTAAGAAAGAGTGGGAAAGATG 120
Db AGAAGCAAGCCCTTCGATGATATTCGCAATATCTCTTAAGAAAGAGTGGGAAAGATG 196
QY 121 AATCTCTCGGAGAAATCGTCTATGTATATGAAGCTAAATATGAGTCTGATGACTAAA 180
Db AATCTCTCGGAGAAATCGTCTATGTATATGAAGCTAAATATGAGTCTGATGACTAAA 256
QY 181 CTAGGTTTCAAGGTCAACCTCCACCTTTCATCGTAGTAGTAAGAGGCTCGACATTCAC 240
Db CTAGGTTTCAAGGTCAACCTCCACCTTTCATCGTAGTAGTAAGAGGCTCGACATTCAC 316
QY 241 GGAATGATTTTGGTAAGATCGAATACCAACACAGGAATCAGTTGAACTCTCAGATGAT 300
Db GGAATGATTTTGGTAAGATCGAATACCAACACAGGAATCAGTTGAACTCTCAGATGAT 376
QY 301 TTGGGAGCTCCAGAGAACTTTCCGAGATCATGCCCAAGAGCCAGCAGAGGAGAA 360
Db TTGGGAGCTCCAGAGAACTTTCCGAGATCATGCCCAAGAGCCAGCAGAGGAGAA 436
QY 361 AATGTTTGAAGAGTGGCAGAGGATCTGGCCCAAAATGATGGGAAAACAGTGTGC 420
Db AATGTTTGAAGAGTGGCAGAGGATCTGGCCCAAAATGATGGGAAAACAGTGTGC 496
QY 421 CCCCAGGAAATCCAGTACCTTGGAGAGATTAACAAGACATCTGCACCCCAAGGGG 480
Db CCCCAGGAAATCCAGTACCTTGGAGAGATTAACAAGACATCTGCACCCCAAGGGG 556
QY 481 AAACATGCTGGACCCACAGACTCGGTGAGAGAAAGCAGCTGTGTTTATGAAGATC 540
Db AAACATGCTGGACCCACAGACTCGGTGAGAGAAAGCAGCTGTGTTTATGAAGATC 616
QY 541 AGCAGCCTGAGGAGATGACGAGTAACTCCCTCG 576
Db AGCAGCCTGAGGAGATGACGAGTAACTCCCTCG 652

RESULT 7
BUI70242
LOCUS BUI70242 873 bp mRNA linear EST 04-SEP-2002
DEFINITION AGENCOURT_7844770 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6051106
5', mRNA sequence.

ACCESSION BUI70242
VERSION BUI70242.1 GI:22684226
KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 873)

AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL COMMENT

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC/DCTD/DRP
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
DNA distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM13304 row: 9 column: 11
High quality sequence stop: 696.

FEATURES

Location/Qualifiers
1..873
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6051106"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC_72"
/note="Organ: skin; Vector: pCMV-SPORT6; Site:1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2 kb. Library constructed by Life
Technologies."

ORIGIN

Query Match 83.6%; Score 481.6; DB 13; Length 873;
Best Local Similarity 89.8%; Pred. No. 3.4e-127;
Matches 517; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 1 ATGAACGGAGACGACCCCTTTCAGAGGAGACCCAGGATGATGCTCAATATCAGAGAAG 60
Db ATGAACGGAGACGACCCCTTTCAGAGGAGACCCAGGATGATGCTCAATATCAGAGAAG 139
QY 61 TTACGAAAGCCCTTCGATGATATTCGCAATATCTCTTAAGAAAGAGTGGGAAAGATG 120
Db AGAAGCAAGCCCTTCGATGATATTCGCAATATCTCTTAAGAAAGAGTGGGAAAGATG 199
QY 121 AATCTCTCGGAGAAATCGTCTATGTATATGAAGCTAAATATGAGTCTGATGACTAAA 180
Db AATCTCTCGGAGAAATCGTCTATGTATATGAAGCTAAATATGAGTCTGATGACTAAA 259
QY 181 CTAGGTTTCAAGGTCAACCTCCACCTTTCATCGTAGTAGTAAGAGGCTCGACATTCAC 240
Db CTAGGTTTCAAGGTCAACCTCCACCTTTCATCGTAGTAGTAAGAGGCTCGACATTCAC 319
QY 241 GGAATGATTTTGGTAAGATCGAATACCAACACAGGAATCAGTTGAACTCTCAGATGACT 300
Db GGAATGATTTTGGTAAGATCGAATACCAACACAGGAATCAGTTGAACTCTCAGATGACT 379
QY 301 TTGGGAGCTCCAGAGAACTTTCCGAGATCATGCCCAAGAGCCAGCAGAGGAGAA 360
Db TTGGGAGCTCCAGAGAACTTTCCGAGATCATGCCCAAGAGCCAGCAGAGGAGAA 439
QY 361 AATGTTTGAAGAGTGGCAGAGGATCTGGCCCAAAATGATGGGAAAACAGTGTGC 420
Db AATGTTTGAAGAGTGGCAGAGGATCTGGCCCAAAATGATGGGAAAACAGTGTGC 499
QY 421 CCCCAGGAAATCCAGTACCTTGGAGAGATTAACAAGACATCTGGAACCCCAAGGGG 480
Db CCCCAGGAAATCCAGTACCTTGGAGAGATTAACAAGACATCTGGAACCCCAAGGGG 559
QY 481 AAACATGCTGGACCCACAGACTCGGTGAGAGAAAGCAGCTGTGTTTATGAAGATC 540
Db AAACATGCTGGACCCACAGACTCGGTGAGAGAAAGCAGCTGTGTTTATGAAGATC 619
QY 541 AGCAGCCTGAGGAGATGACGAGTAACTCCCTCG 576
Db AGCAGCCTGAGGAGATGACGAGTAACTCCCTCG 655

RESULT 8

LOCUS	DEFINITION	888 bp	mRNA	linear	EST 24-MAY-2002
B0432374	AGENCOURT_7859139 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6167925				
5'	mRNA sequence.				
ACCESSION	B0432374				
VERSION	B0432374.1	GI:21171450			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 888)				
AUTHORS	NIH-MGC	http://mgc.nci.nih.gov/.			
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)				
JOURNAL	Unpublished (1999)				
COMMENT	Contact: Robert Strausberg, Ph.D. Email: csapbes-remail.nih.gov Tissue Procurement: ATCC/DCTD/DTF cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LHAM13530 row: j column: 22 High quality sequence stop: 636. Location/Qualifiers 1. 888 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:6167925" /tissue_type="melanotic melanoma" /lab_host="DH10B (phage-resistant)" /clone_lib="NIH MGC 72" /note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 2 kb. Library constructed by Life Technologies."				
FEATURES	source				
ORIGIN					
Query Match	83.6%;	Score 481.6;	DB 13;	Length 888;	
Best Local Similarity	89.8%;	Pred. No. 3.6e-127;			
Matches 517;	Conservative 0;	Mismatches 59;	Indels 0;	Gaps 0;	
QY	1	ATGAACGGAGACGACGCGCTTTCGAAGGAGACCCAGGATGATGCTCAATATCAGAGAG	60		
DB	79	ATGAACGGAGACGACGACCTTTCGAAGGAGACCCAGGATGATGCTCAATATCAGAGAG	138		
QY	61	TTACCAAAAGGCGCTTCGATGATATGCCAAATACCTTCTCTAAGAAAAGATGGGAAAAGATG	120		
DB	139	AGAACAAGGCGCTTTCGATGATATGCCACATCTTCTCTAAGAAAAGATGGGAAAAGATG	198		
QY	121	AAATCCTCGGAGAAAATCGCTATGTGTATATGAAGCTAACTATGAGTCTAGTCTAGCTAAA	180		
DB	199	AAATATCTCGGAGAAAATCAGCTATGTGTATATGAAGAGAACTATAAGGCCATGACTAAA	258		
QY	181	CTAGGTTTCAAGGTCAACCTCCACCTTTCATCGGTAGTAAACGGGCTCGACATTCAC	240		
DB	259	CTAGGTTTCAAGGTCAACCTCCACCTTTCATGTGTATATGAAGCCACACATTCGAG	318		
QY	241	GGGAATGATTTTGGTAAAGATCGAAACCAACAGGAATCAGGTGTAACGTCTCAGATGACT	300		
DB	319	GGGAATGATTTTGGTAAATGACCATACCCAGGATTCAGGTGTAACGTCTCAGATGACT	378		
QY	301	TTCCGACGCTCCAGAGATCTTCCCGAGATCATGCCCAAGAGCCACGAGGAGGAA	360		
DB	379	TTCCGACGCTCCAGAGATCATGCCCAAGATCATGCCCAAGAGCCACGAGGAGGAA	438		
QY	361	AATGTTTGAAGGAAGTGCAGAGGATCTGGCCCAAAATATGATGGGAAAACAGTGTGC	420		
DB	439	AATGTTTGAAGGAAGTGCAGAGGATCTGGCCCAAAATATGATGGGAAAACAGTGTGC	498		

Db 242 CTAGGTTTCAAAGTCACCTCCACCTTTTCATGTGTATTAATAACAGGCCACAGACTTCCAG 301
 QY 241 GGGATGATTTTGGTAACGATCGAACCACAGGATCAGGTTGAACGTCCTCAGATGACT 300
 Db 302 GGAATGATTTTGAATATGACATTAACCGAGGATTCAGGTTGAACATCTCAGATGACT 361
 QY 301 TTCGCGAGCTCCAGAGATCTTCCGAAGATCATGCCCCAAGAGCCAGCAGAGAGAA 360
 Db 362 TTCGCGAGCTCCAGAGATCATCCGAAGATCATGCCCCAAGAGCCAGCAGAGAGAA 421
 QY 361 AATGTTTGAAGAGAGTCCAGAGCATCTGCCACAAATGATGGGAACAGCTGTC 420
 Db 422 AATGATTCGAAGGAGTGCAGAGCATCTGCCACAAACGATGGGAACAACTGCAC 481
 QY 421 CCCCCGGGAAATCCAACTACCTTGGAGAGATTAACAAGACATCTGGACCCCAAGGGG 480
 Db 482 CCCCCAGGAAAGCAATATTTCTGAGAAGATTAATAAGAGATCTGGACCCCAAGGGG 541
 QY 481 AAACATGCTGACCCACAGATCTGCTGAGAGAGCAGCTGCTGTTTATGAGAGATC 540
 Db 542 AAACATGCTGACCCACAGATCTGCTGAGAGAGCAGCTGCTGTTTATGAGAGATC 601
 QY 541 AGCGACCTGAGGAGATCAGGTAACCTCCC 572
 Db 602 AGCGACCTGAGGAGATCAGGTAACCTCCC 633

RESULT 10
 LOCUS CB158955
 DEFINITION K-EST0218399 L18POOL1n1 Homo sapiens cDNA clone L18POOL1n1-32-F07
 5', mRNA sequence.
 ACCESSION CB158955
 VERSION CB158955.1 GI:28145081
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 602)
 Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
 Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
 Kim,Y.S.
 21C Frontier Korean EST Project 2001
 Unpublished (2002)
 Contact: Kim YS
 Genome Research Center
 Korea Research Institute of Bioscience & Biotechnology
 52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
 Tel: +82-42-860-4470
 Fax: +82-42-860-4409
 Email: yongsung@mail.kribb.re.kr
 Plate: 32 row: F column: 07
 High quality sequence stop: 602.

FEATURES

Location/Qualifiers
 1..602
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="L18POOL1n1-32-F07"
 /cell_line="GNU-354-Cho-CK-CK+HLK-3"
 /lab_host="Top10P"
 /clone_lib="L18POOL1n1"
 /note="Organ: Liver; Vector: pT7T3-Pac; Site: 1: EcoRI;
 Site 2: NotI; The library was contributed by the Soares
 Laboratory and it was constructed as described by Bonaldo,
 M.P., Lennon, G. and Soares, M.B. (1996). Genome Research
 6(9): 791-806. RNA was prepared from harvested cell
 culture."

ORIGIN

Query Match 81.9%; Score 471.6; DB 14; Length 602;
 Best Local Similarity 91.1%; Pred. No. 2.2e-124;

Matches 501; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
 QY 1 ATGAACGAGAGACACCGCTTTTGAAGGAGACCCAGGATGATGCTCAATATACAGAGAAG 60
 Db 53 ATGAACGAGAGACACCGCTTTTGAAGGAGACCCAGGATGATGCTCAATATACAGAGAAG 112
 QY 61 TTACGAAAGCCCTTCATGATATTGCAATATCTCTCTAAGAAAGAGTGGGAAAGATG 120
 Db 113 ATCCAAAGAGCCCTTCATGATATTGCAATATCTCTCTAAGAAAGAGTGGGAAAGATG 172
 QY 121 AAATCTCCGAGAAAATCGTCTATGTATATGAAGCTAAACTATAGAGTCACTGACTAAA 180
 Db 173 AAAGCTTCGGAGAAAATCTCTATGTATATGAAGAGAAAGTATGAGGCTATGACTAAA 232
 QY 181 CTAGGTTTCAAGTCAACCTCCACCTTCATGCTAGTAAACGGGCTGCAGACTTCCAC 240
 Db 233 CTAGGTTTCAAGGCCCCTCCACCTTTTCATGTATTAATAACGGGCGGAGACTTCCAG 292
 QY 241 GGAATGATTTTGGTAACGATCGAAACACAGAGAAATCAGGTTGAACGCTCCTCAGATGACT 300
 Db 293 GGAATGATTTGATTAATGACCTTAACCGTGGGAATCAGGTTGAACGCTCCTCAGATGACT 352
 QY 301 TTCGGCAGCTCCAGAGATCTTCCGAAGATCATGCCCAAGAGCCAGCAGAGGAGAA 360
 Db 353 TTGGCAGGCTCCAGGGAATCTCCGAGAGATCATGCCCAAGAGCCAGCAGAGGAGAA 412
 QY 361 AATGTTTGAAGAGAGTCCAGAGGATCTGCCCAACAAATGATGGAAACAGCTGTGC 420
 Db 413 AATGATTCGAGAGAGTCCAGAGAGATCTTGGCCCAAAATGATGGAAAGAGCTGTGC 472
 QY 421 CCCCAGGAAATCCAACTACCTTGGAGAGATTAACAAGACATCTGACCCCAAGAGGGG 480
 Db 473 CCCCAGGAAATCCAACTACCTTGGAGAGATTAACAAGAGATCTGACCCCAAGAGGGG 532
 QY 481 AAACATGCTTGGACCCACAGACTCGCTGAGAGAGAGCAGCTGCTGTTTATGAGAGATC 540
 Db 533 GAACATGCTTGGACCCACAGACTCGCTGAGAGAGAGCAGCTGCTGTTTATGAGAGATC 592
 QY 541 AGCGACCTG 550
 Db 593 AGCGACCTG 602

RESULT 11

BQ231752
 LOCUS BQ231752
 DEFINITION AGENCOURT 7550845 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6066127
 5', mRNA sequence.
 ACCESSION BQ231752
 VERSION BQ231752.1 GI:20413152
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 950)
 NIH-MGC <http://mgs.nci.nih.gov/>
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs@mail.nih.gov
 Tissue Procurement: ATCC/DCTD/DTF
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM1343 row: i column: 08
 High quality sequence stop: 532.
 Location/Qualifiers
 1..950
 /organism="Homo sapiens"

REFERENCE

AUTHORS
 TITLE
 JOURNAL
 COMMENT

FEATURES

source

```
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5066127"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_72"
/notes="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2 kb. Library constructed by Life
Technologies."

ORIGIN
Query Match      81.7%; Score 470.8; DB 13; Length 950;
Best Local Similarity 89.9%; Pred. No. 4.8e-124;
Matches 516; Conservative 0; Mismatches 57; Indels 1; Gaps 1;

QY 1 ATGAACGGAGACGACGCCCTTTGCAAGGAGACCCAGGATGATGCTCAATATACAGAGAG 60
DB 74 ATGAACGGAGACGACGCCCTTTGCAAGGAGACCCAGGATGATGCTCAATATACAGAGAG 133
QY 61 TTACGAAAGCCCTTCGATGATATTGCCAATACTTCTTAAGAAAGATGGGAAAGATG 120
DB 134 ATCAAAAGGCCCTTCGATGATATTGCCAATACTTCTTAAGAAAGATGGGAAAGATG 193
QY 121 AATCTCTCGAGAAATCGCTATGTATGATGATGATGATGATGATGATGATGATGATG 180
DB 194 AAGCCCTCAGAGAAATCTTCTATGTATATGAGAGAAAGATGAGGCTATGACTAA 253
QY 181 CTAGGTTTCAAGGTCACCCCTCCACCTTTTCATCGTGTAGTAAAGGGCTCGACATTCAC 240
DB 254 CTAGGTTTCAAGGTCACCCCTCCACCTTTTCATCGTGTATTAAGCGGCCGAGACTTCCAG 313
QY 241 GGAATGATTTTGGTACGATCGAACCACAGAAATCAGTTGAAAGCTCTCAGATGACT 300
DB 314 GGAATGACGTGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 373
QY 301 TTGGGAGCTCCAGAGAAATCTTCCGAGAGATCATGCCAAGAGCCAGCAGAGAGAGAA 360
DB 374 TTGGGAGCTCCAGGAAATCTTCCGAGAGATCATGCCAAGAGCCAGCAGAGAGAGAA 433
QY 361 AATGTTTGAAGAGTGCAGAGGATCTGGGCCACAAATGATGGGAAACAGCTGTGC 420
DB 434 AATGATTGGAGGAGTGCAGAGGATCTGGGCCACAAATGATGGGAAACAGCTGTGC 493
QY 421 CCCCCGGAAATCAAGTACCTTGGAGAGATTAACAGACATCTGGACCCCAAAAGGGG 480
DB 494 CCCCCGGAAACCAACTACTCTGAGAGATTCACAGAGATCTGGACCCCAAAAGGGG 553
QY 481 AAACAT-GCCTGGACCCACAGACTGCTGAGAGAAAGCAGCTGGTGTATGAGAGAT 539
DB 554 GAAATGCTGGACCCACAGACTGCTGAGAGAAACAGCTGGTGTATGAGAGAT 613
QY 540 CAGGACCCCTGAGGAGATGACGAGTAACCTCC 573
DB 614 CAGGACCCCTGAGGAGATGACGAGTAACCTCC 647

RESULT 12
BI560469
LOCUS 603254435F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5296880 5',
DEFINITION mRNA sequence.
ACCESSION BI560469
VERSION BI560469.1 GI:1544770
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 793)
NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
```

Contact: Robert Strausberg, Ph.D.
Email: cgaabs-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11751 row: e column: 09
High quality sequence stop: 790.

FEATURES
source
1. 793
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5296880"
/lab_host="DH10B"
/clone_lib="NIH_MGC_97"
/notes="Organ: testis; Vector: pBluescriptR (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI
(gtcgag); Oligo-dT primed using primer
5'-TTTTTTTTTTTTTNN-3', size-selected for average
insert size 2.2 kb and normalized to R0T 5. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carninci, in
preparation). Library constructed by M. Brownstein
(NIH/NHGRI, National Institutes of Health). Note: this is
a NIH_MGC Library."

ORIGIN
Query Match 78.9%; Score 454.6; DB 12; Length 793;
Best Local Similarity 90.1%; Pred. No. 2e-119;
Matches 520; Conservative 0; Mismatches 54; Indels 3; Gaps 3;
QY 1 ATGAACGGAGACGACGCCCTTTGCAA-GGAGACCCAGGATGATGCTCAATATACAGAGAA 59
DB 143 ATGAACGGAGACGACGCCCTTTGCAAGGGAGACCCACGGTTGGTCTCAATATACAGAGAA 202
QY 60 GTTACGAAAGCCCTTCGATGATATTGCCAATACTTCTTAAGAAAGATGGGAAAGAT 119
DB 203 GATCCAAAGGCCCTTCGATGATATTGCCAATACTTCTTAAGAAAGATGGGAAAGAT 262
QY 120 GAAATCTCTCGAGAGAAATCGTCTATGTATATGAAGCTAAACTATGAGGTGATGACTAA 179
DB 263 GAAGCCCTCGAGAGAAATCTTCTATGTATATGAAGAGAAATGATGAGGTGATGACTAA 322
QY 180 ACTAGTTTCAAGGTCAACCTCCACCTTCATCGCTAGTAAACGGCTGCAGACTTCCA 239
DB 323 ACTAGTTTCAAGGCCACCCCTCCACCTTTCATGTGTAATAACGGCCGAGAGACTTCCA 382
QY 240 CGGGAATGATTTTGTATACGATCGAAACCCAGAGAAATCAGGTTGAACGTCTCTAGATGAC 299
DB 383 -GGGAATGATTTGGATATGATACCCCTAACCGTGGGAATCAGGTTGAACGTCTCTAGATGAC 441
QY 300 TTTTCGGAGCCTCCAGAGAAATCTTCCCGAAGATCATGCCAAGAGCCAGCAGAGAGAGA 359
DB 442 TTTTCGGAGCCTCCAGAGAAATCTTCCCGAAGATCATGCCAAGAGCCAGCAGAGAGAGG 501
QY 360 AAATGTTTGAAGGAAGTGCAGAGGATCTTGGCCCAAAATGATGGGAAACAGCTGTG 419
DB 502 AAATGATTCGGAGGAGTGCAGAGGATCTTGGCCCAAAATGATGGGAAAGAGCTGTG 561
QY 420 CCCCCGGGAAATCAAGTACCTTGGAGAGATTAACAGACATCTGAGACCCAAAGGGG 479
DB 562 CCCCCGGGAAACCAACTACTCTGAGAGATTCAGAGAGATCTGAGACCCAAAGGGG 621
QY 480 GAAACATGCTCGAGCCACAGACTGCGTGAGAGAAAGCAGCTGGTGTATGAA-GAGA 538
DB 622 GGAACATGCTCGAGCCACCCAGACTGCGTGAGAGAAACAGCTGGTGTATGAAACGAGA 681
QY 539 TCAGCGACCCCTGAGGAGATGACGAGTAACCTCCCTC 575


```
|||||
682 TCAGGACCTGAGAAAGATGACGAGTAACCTCCCTC 718

RESULT 13
BQ229996      922 bp      mRNA      linear      EST 02-MAY-2002
AGENCOURT_7560406 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6048424
5', mRNA sequence.
ACCESSION    BQ229996
VERSION      BQ229996.1 GI:20411396
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC/DCTD/DTP
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13297 row: g column: 17
High quality sequence stop: 483.
Location/Qualifiers
1..922
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6048424"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 72"
/notes="Organ: Skin; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2 kb. Library constructed by Life
Technologies."

FEATURES
Source
Query Match 77.6%; Score 447; DB 13; Length 922;
Best Local Similarity 89.1%; Pred. No. 3.2e-117; Indels 3; Gaps 3;
Matches 516; Conservative 0; Mismatches 60;

QY 1 ATGAACGGAGACGACGCTTTGCAAGAGACCCAGGAGTATGCTCAAAATATCAGAGAAG 60
DB 44 ATGAACGGAGACGACACCTTTGCAAGAGACCCAGGAGTATGCTCAAAATATCAGAGAAG 103
QY 61 TTAGGAAGCCCTCGATGATATGCAATATCTCTTAAGAAAGAGTGGGAAGAAGT 120
DB 104 AGAAGCAAGCCCTTGATGATATGCCCACATCTTCTTAAGAAAGTGGAAAGAAGT 163
QY 121 AATCTCTCGAGAAATCGTCTATGTATATGAAGCTAACTATGAGTCTAGCTATAA 180
DB 164 AATACTCGAGAAATCAGCTATGTATATGAAGAAACTATTAAGGCCATGACTAAA 223
QY 181 CTAGTTTCAAGTCAACCTCCACCTTTTCATCGGTAGTAAACGGGTGACAGCTCCAC 240
DB 224 CTAGTTTCAAGTCAACCTCCACCTTTTCATCGGTAGTAAACGGGTGACAGCTCCAC 283
QY 241 GGGATGATTTTGTATACGATCGAAACACAGGAATCAGGTTGAACTCTCTCAGATGACT 300
DB 284 GGGATGATTTTGTATACGATCGAAACACAGGAATCAGGTTGAACTCTCTCAGATGACT 343
QY 301 TTCGGCAGCCTCCAGAGAATCTTCCGAGAGATCATGCCCAAGAGCCAGCAGAGGAAGAA 360
DB 344 TTCGGCAGCCTCCAGAGAATCTTCCGAGAGATCATGCCCAAGAGCCAGCAGAGGAAGAA 403
```

```
|||||
361 AATGTTTGAAGAGTCCAGAGGATCTGCCCCACAAATGATGGGAAACAGCTGTGC 420
DB 404 AATGATTCGAAGGGAGTGTCAAGAGCATCTGCCCCACAAACGATGGGAAACACTGCAC 463
QY 421 CCCCCGGGAAATCCAAAGTACCTTGGAGAAAGATTAAACAGACATCTGGACCCCAAAA- 479
DB 464 CCCCCGAGAAACCAATTAATCTTGAGAGATTAAATAGAGATCTGGACCCCAAGGGGG 523
QY 480 GAAACATGCTGACCCACAGACTGCTGAGAGAAAGACAGCTGGTGG- 538
DB 524 GAAACATGCTGACCCACAGACTGCTGAGAGAAAGACAGCTGGGGGATTTATGAAGAGA 583
QY 539 TCAG-CGACCTCGAGGAAGATGACGAGTAACCTCCCTCG 576
DB 584 TCAGCCGACCTCGAGGAAGATGACGAGTAACCTCCCTCG 622

RESULT 14
BE408883      573 bp      mRNA      linear      EST 21-JUL-2000
LOCUS        601303758F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:3637982 5',
DEFINITION   mRNA sequence.
ACCESSION    BE408883
VERSION      BE408883.1 GI:9345333
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCM338 row: d column: 15
High quality sequence stop: 571.
Location/Qualifiers
1..573
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3637982"
/tissue_type="choriocarcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 21"
/notes="Organ: Placenta; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCAGGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

```
Query Match 73.6%; Score 423.8; DB 10; Length 573;
Best Local Similarity 88.9%; Pred. No. 1.2e-110;
Matches 458; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY 1 ATGAACGGAGACGACGCTTTGCAAGAGACCCAGGAGTATGCTCAAAATATCAGAGAAG 60
DB 59 ATGAACGGAGACGACACCTTTGCAAGAGACCCAGGAGTATGCTCAAAATATCAGAGAAG 118
QY 61 TTACGAAGGCTTCGATGATATTCGCAATACTTCTTAAGAAAGAGTGGAAAGATG 120
```

Db 119 AGAAGCAAGGCTTTGATGATATTGCCACATCTTCTTAAGAAAGAGTGGAAAAAGATG 178
Qy 121 AATCTCTCGAGAAATCGCTATGTATATGATGATGATGATGATGATGATGATGATGATGAT 180
Db 179 AATATCTCGAGAAATCGCTATGTATATGATGATGATGATGATGATGATGATGATGATGAT 238
Qy 181 CTAGGTTTCAAGTCAACCTCCACCTTTCATGCGTAGTAAAGCGGCTCCAGACTTCCAC 240
Db 239 CTAGGTTTCAAGTCAACCTCCACCTTTCATGCGTAGTAAAGCGGCTCCAGACTTCCAG 298
Qy 241 GGAATGATTTTGGTACGATCGAACCACAGGAATCAGGTTGAAGCTCCTCAGATGACT 300
Db 299 GGAATGATTTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 358
Qy 301 TTCCGCGAGCTCCAGAGAAATCTTCCGGAAGATCATGCCCAAGAGCCAGCAGAGGAAGAA 360
Db 359 TTCCGCGAGCTCCAGAGAAATCATCTCCGGAAGATCATGCCCAAGAGCCAGCAGAGGAAGAA 418
Qy 361 AATGTTTGAAGAGTCCAGAGGATCTGGCCCAACAAATGATGGGAAACAGCTGTGC 420
Db 419 AATGATTCGAAGGAGTGTGAGAGCATCTGGCCCAACAAATGATGGGAAACAGCTGTGC 478
Qy 421 CCCCCGGGAAATCCAAAGTCTTGGAGAGATTAAACAAGACATCTGGACCCCAAAAGGGGG 480
Db 479 CCCCCGGGAAAGCAATATTTCTGAGAGATTAAACAAGATCTGGACCCCAAAAGGGGG 538
Qy 481 AAACATCCTCGAGCCACAGACTGCGTGAGAGAAA 515
Db 539 AAACATCCTCGAGCCACAGACTGCGTGAGAGAAA 573

RESULT 15
BE891434
LOCUS 601434893f1 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:3919991 5',
- DEFINITION mRNA sequence.
ACCESSION BE891434
VERSION BE891434.1 GI:10350764
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE NIH-MGC http://mgi.nci.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC/DCTD/DTF
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM9750 row: j column: 24
High quality sequence stop: 641.
Location/Qualifiers
1..867
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3919991"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_72"
/note="Organ: skin; Vector: pCMV-SPORT6; Site: 1; NotI;
Site 2: SalI; Cloned unidirectionally. Primer: oligo dt.
Average insert size 2 kb. Library constructed by Life
Technologies."

FEATURES
source

ORIGIN

Query Match 73.5%; Score 423.4; DB 10; Length 867;
Best Local Similarity 89.4%; Pred. No. 1.9e-110;
Matches 512; Conservative 0; Mismatches 56; Indels 5; Gaps 5;
Qy 1 ATGAAACGGAGACGACGCTTTGCAAGGAGACCCAGGAGATGATGCTCAAAATATCAGAGAAG 60
Db 50 ATGAAACGGAGACGACGCTTTGCAAGGAGACCCAGGAGATGATGCTCAAAATATCAGAGAAG 109
Qy 61 TTACGAAAGGCTTTGATGATATTGCGCAAAATACCTTCTTAAGAAAGAGTGGGAAAAAGATG 120
Db 110 ATCCAAAGGCTTTGATGATATTGCGCAAAATACCTTCTTAAGGAGAGTGGGAAAAAGATG 169
Qy 121 AATTCCTCGAGAAATCGTCTATGTATATGAGCTAACTATGAGTCTAGCTACTATAA 180
Db 170 AAGCCCTCAGAGAAATCTTCTATGTGTATATGAGAGAAAGTATGAGGCTATGACTTAA 229
Qy 181 CTAGGTTTCAAGGCTCCAGAGAAATCTTCCGGAAGATCATGCCCAAGAGCCAGCAGACTTCCAC 240
Db 230 CTAGGTTTCAAGGCTCCAGAGAAATCTTCCGGAAGATCATGCCCAAGAGCCAGCAGACTTCCAC 288
Qy 241 GGAATGATTTTGGTAAAGATCGAACCACAGGATCAGGTTGACGCTTGNACGCTCTCAGATGACT 300
Db 289 GGAATGATTTTGGTAAAGATCGAACCACAGGATCAGGTTGACGCTTGNACGCTCTCAGATGACT 348
Qy 301 TTCCGCGAGCTCCAGAGAAATCTTCCGGAAGATCATGCCCAAGAGCCAGCAGAGGAAGAA 360
Db 349 TTCCGCGAGCTCCAGAGAAATCTTCCGGAAGATCATGCCCAAGAGCCAGCAGAGGAAGAA 408
Qy 361 AATGTTTGAAGAGTCCAGAGGATCTGGCCCAACAAATGATGGGAAACAGCTGTGC 420
Db 409 AATGATTCGAAGGAGTCCAGAGGATCTGGCCCAACAAATGATGGGAAACAGCTGTGC 468
Qy 421 CCCCCGGGAAATCCAAAGTCTTGGAGAGATTAAACAAGACATCTGGACCC-AAAAAGGGG 479
Db 459 CCCCCGGGAAATCCAAAGTCTTGGAGAGATTAAACAAGACATCTGGAGAGATCTGGACCCAAAGGGG 527
Qy 480 GAAACATGCTCGAGCCCAAGAGATGCGGTGAGAGAAAGCAGCTGGTGGTTTATGAGAGAT 539
Db 528 GGAACATGCTCGAGCCCAAGAGATGCGGTGAGAGAAAGCAGCTGGTGGTTTATGAGAGAT 586
Qy 540 CAGCGACCTCAGGAGATGACGAGTAACCTCC 572
Db 587 CAGCGA-CCTGAGGAGATGACGAGTAACCTCC 618

Search completed: March 31, 2004, 17:25:05
Job time : 2750 secs